Performance of the LHCb Vertex Locator and flavour tagging studies for the measurement of the CKM angle y

Farinelli, C.

2014

document version
Publisher's PDF, also known as Version of record

Introduction

Subatomic particles and their interactions are described by the Standard Model (SM) [1], which has proven to be a successful theory as its predictions have been confirmed by experimental observations ever since its conception forty years ago. However, the SM cannot be considered to be complete since it does not include a theory for gravitation and offers no candidate for dark matter, which makes up about 85% of all matter in the universe. In addition, although the SM allows for matter-antimatter asymmetries in particle interactions, generally referred to as CP violation, the amount of CP violation that is predicted cannot explain the dominance of matter over antimatter in our universe.

Overview of the Standard Model

The fundamental elementary particles are classified into two groups: fermions, with half-integer spin, and bosons, whose spin is an integer. The elementary matter constituents (i.e. quarks and leptons) are fermions, which are subdivided into three generations having similar properties but different masses, as illustrated in Fig. 1.

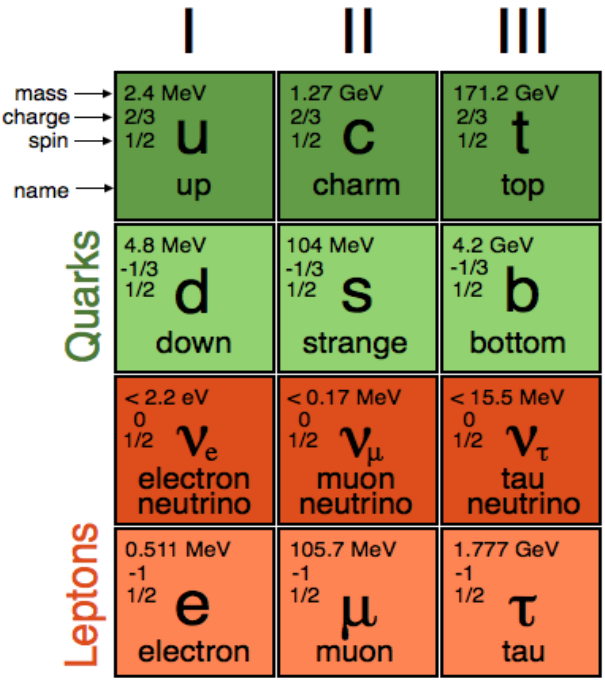

Figure 1: *The fermions in the Standard Model: quarks and leptons. For each fermion, the values of the mass, charge and spin are shown.*

Interactions between fermions occur via forces carried by bosons: the strong force, the weak

force and the electromagnetic force. In Table 1, the various bosons of the SM are listed, together with the associated forces and their relative strength.

Force	Charge	Boson	Relative strength
strong	colour	gluon (g)	$O(1)$
electromagnetic	electric charge	photon (γ)	$O(10^{-2})$
weak	weak isospin	Z^0, $W^{\pm}$	$O(10^{-5})$

Table 1: *Bosons in the Standard Model. The force carried by each boson and the charge on which it acts are listed, together with their relative coupling strengths at low momentum transfer.*

The electromagnetic force acts on all charged particles through the exchange of a photon, the weak force acts on all fermions and is mediated by the $W^{\pm}$ and Z^0 bosons and the strong force, carried by eight flavours of gluons, acts only on the colour quantum number (*red, blue, green*) of quarks.

Due to the nature of the weak interaction each generation of quarks contains an *up-type* and a *down-type* quark, while each generation of leptons contains a charged lepton ($e^{\pm}$, $\mu^{\pm}$, $\tau^{\pm}$) and a corresponding (neutral) neutrino. Similarly, the strong interaction leads to three different colour states for each type of quark (not indicated in the figure).

Finally, in addition to the spin-1/2 fermions and spin-1 bosons the SM includes a spin-0 particle known as the Higgs boson, for which recent experimental evidence was seen [2, 3].

In the Standard Model to each particle with non-zero internal quantum numbers (e.g. charge) an antiparticle with opposite charge is associated. Composite neutral particles (e.g. the π^0) can be their own antiparticle as well as the neutral force carriers γ and Z^0. It remains an open question whether the neutrino is its own antiparticle or not. Due to the theorem of CPT invariance [4, 5, 6], the mass and decay rate of particles and corresponding antiparticles must be equal.

In nature, quarks can only exist in bound states known as hadrons. There are two types of hadrons: baryons, that are made up of three quarks or three antiquarks, and mesons, made up of a quark and an antiquark. The work of this thesis focuses on B_s mesons, which are made up by a $b(\bar{b})$-quark and an $\bar{s}(s)$-quark.

Symmetries in the SM and CP violation

The Standard Model is based on symmetries and conservation laws, and can be summarized in group theory by the structure:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$$

In this expression, the first term represents the theory of strong interactions, known as Quantum Chromodynamics (QCD), while the second and third term describe the unified electroweak force.

Some of the symmetries of the SM are believed to be exact, while others are only approximately valid; in the latter case a symmetry is said to be broken or violated. In case of exact symmetries the quanta of the corresponding gauge fields (the force carriers) are predicted to be massless.

The Brout-Englert-Higgs mechanism [7] provides a description of how the force carriers of the weak interaction, the $W^{\pm}$ and Z^0, receive a mass via the spontaneous symmetry breaking of the electroweak symmetry, whereas the photon remains massless.

Together with these continuous gauge symmetries, alternative symmetries under discrete operations of space inversion (Parity P), charge inversion (C) and time reversal (T) can be considered. The strong interaction is believed to be symmetric under all these three symmetries

individually. The broken electroweak symmetry results in electromagnetism, which is also symmetric under C, P and T operations individually, and the weak interaction which is not. In fact, the weak interaction was seen to maximally violate the left-right Parity symmetry.

The weak interaction was then thought to be symmetric under the simultaneous operation of parity and charge inversion, and no violation of the CP symmetry had been observed until 1964, when J. Cronin and V. Fitch observed a small amount of CP violation in the neutral kaon system [8][1]. At the beginning of this century, proof of CP violation in the B meson system was observed first at the BaBar and Belle experiments [9, 10] and, subsequently, at the D0 and CDF experiments at the Tevatron accelerator complex at Fermilab [11, 12]. Since the SM predicts large CP violation effects in the B meson sector, these particles are particularly suited to study the violation of this symmetry.

In the Standard Model flavour changing interactions between quarks are mediated by the $W^{\pm}$ boson. At the same time the quarks couple to the Higgs field by means of so-called Yukawa couplings, giving rise to mass terms. The flavour eigenstates of the quark fields are, however, not identical to the mass eigenstates and their transformation includes complex phases. These complex phases give rise to CP violation in the flavour changing quark decays. The framework of the flavour changing quark interactions is referred to as the Cabibbo-Kobayashi-Maskawa (CKM) mixing mechanism [13, 14], and will be described further in Chapter 1 of this thesis.

Using the unitarity of the CKM theory, a single complex phase in the Yukawa couplings translates into triangle relations in the complex plane of charged current couplings, leading to 6 unitarity triangles with angles α, β and γ. The main research question of quark flavour physics is whether all CP violating processes are generated by this mechanism, which has its origin in the electroweak phase transition.

In particular, the unitarity angle γ is confronted with consistency tests of the CKM picture. Measurements of the angle γ using time-dependent analyses of neutral B_d^0 and B_s^0 mesons are possible and rely on the interference of $b \to u$ and $b \to c$ amplitudes achieved through B_d^0 or B_s^0 mixing. The tree-level decays $B_s^0 \to D_s^{\mp} K^{\pm}$, on which the analysis presented in this thesis is performed, are particularly suited for these measurements, since the amplitudes of the decays $B_s^0 \to D_s^- K^+$ and $B_s^0 \to D_s^+ K^-$ have comparable magnitudes and therefore the interference effects are expected to be large.

B-physics at the LHCb experiment

The LHCb experiment at the Large Hadron Collider (LHC) was designed to study beauty and charm flavour physics, and currently plays a leading role in measurements of CP violation. At present the LHC is the most abundant source of b quarks, which can hadronize into different types of long lived B mesons and baryons. B hadrons are produced in pairs that fly predominantly in the forward or backward direction of LHCb. These properties lead to the design of a forward spectrometer and make LHCb optimal for the study of CP violation in the B-meson sector, as well as for the search for rare decays of B mesons.

In addition to accommodate for the large $b\bar{b}$ production cross section, the LHCb spectrometer was optimized to do flavour physics precision studies. An excellent vertex resolution is needed in order to disentangle secondary vertices produced by B mesons from primary vertices produced from pp interactions. To measure the oscillations of B mesons the measurement of the proper time must be precise, especially for measurements of the fast-oscillating B_s mesons. The *hardware* research of this thesis focuses on the VErtex LOcator (VELO) subdetector, used for vertexing at LHCb. Moreover, a good momentum and invariant mass resolution as well as clean particle identification are needed to separate b hadrons from the abundant backgrounds,

[1]The conservation of the CP-symmetry in strong interactions is known as the *strong CP problem* and consists in the fact that, although there is no reason why the CP-symmetry should be conserved in QCD specifically, no violation of CP has ever been observed experimentally in strong interactions.

as in any flavour physics experiment.

Overview of dissertation

The work described in this thesis can be divided into two main parts: the first involves commissioning work and data quality analyses for the VELO detector, and the second focuses on the study of CP violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays with emphasis on flavour tagging.

In Chapter 1 the theory of weak interactions and how CP violation is incorporated in the SM are discussed. Also, a description of how CP observables can be extracted from time-dependent B_s^0 decays is given. A description of the different subdetectors of LHCb and their purpose is provided in Chapter 2. In Chapter 3 the VELO detector is presented. The geometry and data acquisition system of the VELO are described, and possible errors that can occur during the readout of data are listed. Chapter 4 and Chapter 5 are dedicated to the study of erroneously encoded hits in the VELO detector and the corresponding signals that are transmitted to the readout channels; in Chapter 4 channels having high ADC signals are analysed, while in Chapter 5 negative signals are studied.

The analysis on $B_s^0 \to D_s^{\mp} K^{\pm}$ decays is presented in Chapters 6, 7 and 8, aiming at increasing the precision of the measurement of the CKM angle γ. For these precision measurements it is important to correctly identify the flavour of the B_s^0 meson at production, and in particular determine the fraction of B_s^0 candidates that have been misidentified. The calibration of this *mistag fraction* is performed on the topologically similar $B_s^0 \to D_s^- \pi^+$ events, since for this channel the flavour of the B_s^0 at decay can be unequivocally determined by identifying the charge of the final state pion. In Chapters 6 and 7, the calibration of the mistag fraction using $B_s^0 \to D_s^- \pi^+$ events collected at LHCb during 2010 and 2011 is treated. Finally, sensitivity studies for the measurement of the CKM angle γ are performed on generated $B_s^0 \to D_s^{\mp} K^{\pm}$ events. These studies are presented in Chapter 8.

Chapter 1

CP violation in the B_s^0 meson system

The physics of interactions between different quark types is referred to as flavour physics. The rich structure of flavour physics is for a large extent due to the quark mixing mechanism as proposed by Cabibbo [13], Kobayashi and Maskawa [14].

In Sec. 1.1 of this chapter we review the origin of quark generation mixing in the Standard Model, which is due to the Yukawa couplings of the Higgs field to quarks. We will see in Sec. 1.2 how this leads to flavour changing interactions that allow for CP-violating processes to occur. The amount of CP violation in the theory is represented by the unitarity triangle, which is characterized by the observable angles α, β and γ. The neutral B-$\overline{B}$ system[1], discussed in Sec. 1.3, is well suited to study CP violating processes, since decays of particles containing a b-quark are expected to lead to relatively large CP violating observables.

In general, the occurrence of CP violation can be classified in three different types. The different types of CP violation are introduced in Sec. 1.4, together with a description of how they can be measured in the B meson system. The specific decay $B_s^0 \to D_s^{\mp} K^{\pm}$ is worked out in Sec. 1.5 and allows for a determination of the angle γ in a method that is free of theoretical uncertainties. Since this decay does not have large contributions from loop diagrams it is not likely affected by possible new particles and serves as a reference γ measurement in the CKM framework.

Finally in Sec. 1.6 we illustrate the experimental signature of the CP-violating decay rates and explain the main measurement dilution effects. One of these dilutions, the flavour tagging, is discussed in more detail in Chapter 6 of this thesis.

1.1 Weak Interactions and the CKM Matrix

The Standard Model (SM) describes strong, weak and electromagnetic interactions of fundamental particles according to the principle of gauge invariance [15, 16]. This means that the theory is based on certain symmetries of nature that, according to Noether's theorem [17], in general result in conserved quantities. The strong interaction is described by the exact $SU(3)_C$ symmetry group, where C represents the colour quantum number of quarks. The electroweak interactions are based on the $SU(2)_L \otimes U(1)_Y$ symmetry group, where L is the left-handed weak isospin quantum number and Y is the hypercharge, defined in terms of the electric charge Q and the third component of the isospin I_3 as $Y = 2(Q - I_3)$ [18, 19, 20]. This electroweak symmetry is spontaneously broken by the Higgs mechanism [7]. The spontaneous symmetry breaking generates masses for particles and separates the electromagnetic and weak forces mediated by, on the one hand, the massless photon γ and, on the other hand, the massive $W^{\pm}$ and Z^0 bosons, respectively.

[1] with neutral B mesons we refer to both B_d^0 and B_s^0 mesons.

The flavour changing interactions between quarks take place via the charged currents carried by the W^+, W^- bosons:

$$\mathcal{L}_{cc} = \frac{g}{\sqrt{2}}\bar{u}_L^i \gamma_\mu W^{+\mu} d_L^i + \frac{g}{\sqrt{2}}\bar{d}_L^i \gamma_\mu W^{-\mu} u_L^i + h.c. \tag{1.1}$$

where g is the weak coupling constant, γ_μ are the Dirac matrices ($\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$), and $u_L^i = \left(\frac{1-\gamma_5}{2}\right) u^i$ projects out the left-handed spinor component. The quark spinor fields include $u^i = (u, c, t)$ and $d^i = (d, s, b)$.

In the SM, the coupling between the Higgs field ϕ and the fermion fields is described by the Yukawa interaction, which is of the form:

$$\mathcal{L}_{Yukawa} = Y_{i,j}^u (\bar{Q}_L^i \tilde{\phi}) u_R^j + Y_{i,j}^d (\bar{Q}_L^i \phi) d_R^j + h.c. \tag{1.2}$$

In this expression, the Higgs field ϕ is a complex spinor doublet:

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$$

and $\tilde{\phi} = i\tau_2\phi$, where τ_2 is the second Pauli matrix. The fields $Q_L = (u_L \; d_L)$ are the left-handed weak interaction doublets and u_R, d_R are right-handed singlets. In the interaction representation of the quark fields the charged current (Eq. 1.1) is flavour diagonal while the Yukawa coupling (Eq. 1.2) is not. The Yukawa matrices $Y^{u,d}$ are 3×3 complex matrices for the three generations of quarks (i, j). Alternatively the expressions can be written using the mass eigenstates of the quark fields, by diagonalizing the Yukawa matrices via unitary transformations $V_{R,L}^{u,d}$, representing rotations of the fields.

After electroweak symmetry breaking, the Higgs field condensates in a ground state of non-zero vacuum expectation $v = \langle \phi^0 \rangle$ and the resulting mass terms for quarks are of the form:

$$M^u = \frac{v}{\sqrt{2}} V_L^u Y^u V_R^{u\dagger} \quad , \quad M^d = \frac{v}{\sqrt{2}} V_L^d Y^d V_R^{d\dagger}$$

where the Yukawa matrices have been diagonalized with a bi-unitary transformation $Y^{u,d} \rightarrow V_L^{u,d} Y^{u,d} V_R^{u,d\dagger}$. The charged current interaction terms in the Lagrangian can then be written in the same mass representation of the quark fields $u_L = V_L^{u\dagger} u_L^i$ and $d_L = V_L^{d\dagger} d_L^i$, as:

$$\mathcal{L}_{cc} = \frac{g}{\sqrt{2}} \left(\begin{bmatrix} \bar{u}_L & \bar{c}_L & \bar{t}_L \end{bmatrix} \gamma_\mu W^{+\mu} V_{CKM} \begin{bmatrix} d_L \\ s_L \\ b_L \end{bmatrix} + \begin{bmatrix} \bar{d}_L & \bar{s}_L & \bar{b}_L \end{bmatrix} \gamma_\mu W^{-\mu} V_{CKM}^\dagger \begin{bmatrix} u_L \\ c_L \\ t_L \end{bmatrix} \right) \tag{1.3}$$

where V_{CKM} is now the complex unitary matrix referred to as CKM (Cabibbo-Kobayashi-Maskawa) matrix [13, 14]. The CKM matrix elements are given as:

$$(V_{CKM})_{ij} = (V_L^u V_L^{d\dagger})_{ij},$$

and determine the strength of the interactions between quarks of different flavour and the $W^\pm$ boson. Its elements, when multiplied by $g/\sqrt{2}$, give the coupling of the weak current and need to be determined experimentally. The unitarity of the CKM matrix accounts for the absence of flavour changing neutral current (FCNC) processes at tree level in the SM and represents a generalization of the GIM (Glashow-Ilioupoulos-Maiani) mechanism [21] to three quark generations.

By convention, the interaction and mass eigenstates are taken the same for *up*-type quarks but are rotated for *down*-type quarks, resulting in:

$$u_i = u_j^{Mass}$$
$$d_i = (V_{CKM})_{ij} d_j^{Mass}$$

From Eq. 1.1 it follows that the SM Lagrangian is invariant under CP transformations (i.e. $u \to \bar{u}$, $d \to \bar{d}$, $W^+ \to W^-$, $W^- \to W^+$, $V_{CKM} \to V_{CKM}^\dagger$) only if $V_{ij} = V_{ij}^*$, implying that a non-vanishing complex phase in the CKM matrix can generate CP violating transitions.

The existence of a third generation of quarks was indeed predicted by Kobayashi and Maskawa in order to allow for the existence of an imaginary degree of freedom in the CKM matrix to explain CP violating phenomena. In general, an $n \times n$ complex matrix has $2n^2$ free parameters. Since V_{CKM} is a unitary matrix, the degrees of freedom are reduced by n^2, leaving us with n^2 free parameters. In addition, the relative phases of the $2n-1$ quark fields can be chosen, resulting in the CKM matrix having $n^2 - (2n-1) = (n-1)^2$ degrees of freedom.

For an orthogonal $n \times n$ matrix there are, in general, $\frac{1}{2}n(n-1)$ independent rotation angles. This means that, out of the original $(n-1)^2$ free parameters $\frac{1}{2}n(n-1)$ are rotation angles and $\frac{1}{2}(n-1)(n-2)$ are independent complex phases [22]. Therefore, in order to allow for the presence of a complex phase for CP violation there must be at least $n = 3$ generations, such that the CKM matrix can be described by four physical parameters: three rotation angles and one complex phase.

1.2 The CKM Matrix and Unitarity Triangle

All of the CP-violating effects in the flavour sector of the SM originate from the complex phase of the CKM matrix. The CKM matrix can be written as:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.4}$$

The unitarity requirement of the matrix V implies 3 normalizations:

$$\sum_{k=1}^{3} V_{ik} \cdot V_{ik}^* = 1 \qquad \forall i \in 1,2,3 \tag{1.5}$$

and 6 orthogonality relations:

$$\sum_{k=1}^{3} V_{ik}^* \cdot V_{jk} = 0 \qquad \forall i,j \in 1,2,3 (i \neq j) \tag{1.6}$$

$$\sum_{k=1}^{3} V_{ki} \cdot V_{kj}^* = 0 \qquad \forall i,j \in 1,2,3 (i \neq j) \tag{1.7}$$

The orthogonality relations identify 6 triangles in the complex plane. These triangles all have the same surface area, which can be interpreted as a measure of the CP violation in the SM [23].

For phenomenological applications it is useful to parametrize the CKM matrix according to the Wolfenstein parametrization [24], which reflects the measured hierarchy between the elements of the CKM matrix, as:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \tag{1.8}$$

where $A \simeq 0.81$, λ=$\sin\theta_C \simeq 0.22$ (θ_C is the Cabibbo angle) gives the hierarchy of the coupling strengths, and ρ and η parametrize the amount of CP violation in the SM[2].

Using the Wolfenstein parametrization it is straightforward to observe that four of the six triangles originating from the orthogonality relations are squashed (one side being much shorter than the other two), one of them is somewhat squashed (i.e. the "B_s triangle"), while for one of them (the so-called B_d triangle), all sides have approximately the same length. The relation describing the B_d triangle is:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.9}$$

where each of the terms is of the order λ^3.

The B_d triangle is usually referred to as the Unitary Triangle (UT) of the CKM matrix. The UT can be normalized by dividing the previous expression by $V_{cd}V_{cb}^*$, such that one of the sides becomes of unit length:

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} + 1 = 0 \tag{1.10}$$

The lengths of the sides of the UT are then given by:

$$R_c = 1 \tag{1.11}$$

$$\begin{aligned} R_u = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| &= \left|\left(1-\frac{\lambda^2}{2}\right)(\rho+i\eta)\right| \\ &= \left(1-\frac{\lambda^2}{2}\right)\sqrt{\rho^2+\eta^2} \equiv \sqrt{\bar{\rho}^2+\bar{\eta}^2} \end{aligned} \tag{1.12}$$

$$R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = |1-\rho-i\eta| \equiv \sqrt{(1-\bar{\rho})^2+\bar{\eta}^2} \tag{1.13}$$

where $(\bar{\rho},\bar{\eta})$ are the coordinates in the complex plane of the only non-trivial apex of the UT (the others are $(0,0)$ and (1,0)).

The CKM matrix has complex elements, but the individual phases do not have physical meaning, as described in Sec. 1.1. Physical meaningful quantities must be independent of quark rephasing; the simplest ones are the magnitudes of the CKM elements $|V_{ij}|^2$. The next simplest ones that do not depend on rephasing are combinations of four elements: $V_{\alpha i}V_{\beta j}V_{\alpha j}^*V_{\beta i}^*$, with $\alpha \neq \beta$ and $i \neq j$. Examples are the internal angles of the B_d triangle; α, β, γ:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \qquad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \qquad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{1.14}$$

In the Wolfenstein parametrization to $O(\lambda^3)$, $-\beta$ is the phase in the matrix element V_{td} and $-\gamma$ represents the phase in the element V_{ub}.

The triangle corresponding to the B_s meson is given by the orthogonality relation:

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.15}$$

In this expression the d-quark in Eq. 1.9 has been replaced with the s-quark. The angles of the B_s triangle α_s, β_s and γ_s are given by:

$$\alpha_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{us}V_{ub}^*}\right) \qquad \beta_s = \arg\left(-\frac{V_{cs}V_{cb}^*}{V_{ts}V_{tb}^*}\right) \qquad \gamma_s = \arg\left(-\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*}\right) \tag{1.16}$$

[2]CP-violation implies $\eta \neq 0$.

and the value of β_s is approximately one degree [25].

In Fig. 1.1, the unitary triangle in the B_d system (see Eq. 1.9) is depicted on the left. On the right, the unitary triangle in the B_s system (see Eq. 1.15) is shown.

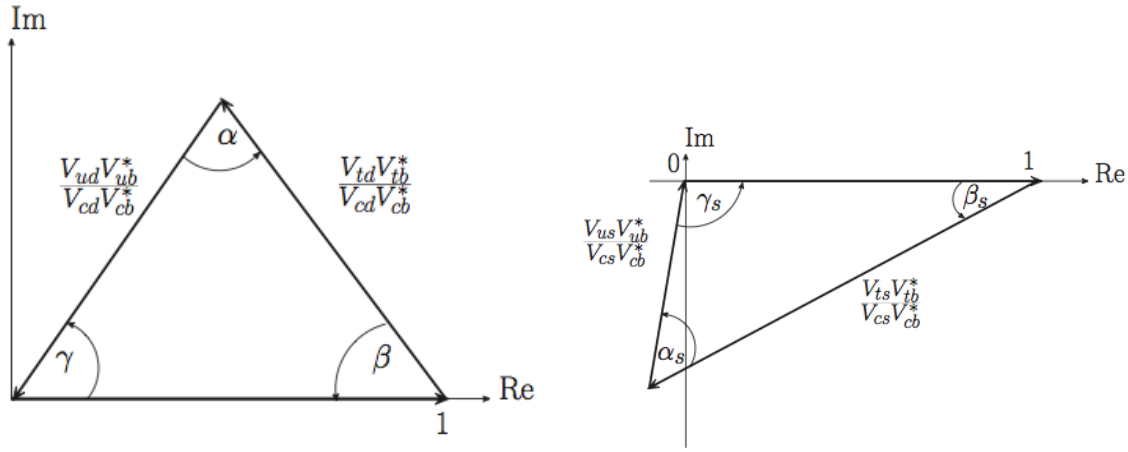

Figure 1.1: *The normalized unitary triangle in the B_d system (left) and B_s system (right).*

One of the goals of B-physics is to experimentally measure and overconstrain the angles and sides of the B_d and B_s triangles to verify the validity of the SM predictions, which for all flavour changing interactions should be described by the CKM mechanism. Before the start of the LHC, the highest precision measurements on the B_d system had been performed at the B-factories PEP-II by the BaBar [78] collaboration and at KEK-B by the Belle [79] collaboration, while the measurements on the B_s system had been performed at the Tevatron by the CDF [80] and D0 [81] collaborations. In the meantime LHCb has become the leading facility.

B-physics measurements are generally summarized by allowed regions in the $(\bar{\rho},\bar{\eta})$ plane, where $\bar{\rho}_{(s)}$ and $\bar{\eta}_{(s)}$ are defined as:

$$\bar{\rho} = \left(1 - \frac{\lambda^2}{2}\right)\rho \quad , \quad \bar{\eta} = \left(1 - \frac{\lambda^2}{2}\right)\eta$$
$$\bar{\rho}_s = \left(\frac{-\lambda^2}{1-\lambda^2/2}\right)\rho \quad , \quad \bar{\eta}_s = \left(\frac{-\lambda^2}{1-\lambda^2/2}\right)\eta$$

and give the apex of the B_d and B_s triangles respectively [26].

Fig. 1.2 shows the current constraints on the location of the apex of the B_d unitarity triangle (UT) and the B_s unitarity triangle (UT_s). From the figure it can be seen that currently all measurements are consistent[3] with the CKM description of weak interactions in the SM.

The validity of the SM picture can be further tested by comparing UT observables resulting from tree-diagram dominated decays to the same observables resulting from loop-diagram dominated decays (e.g. penguins). In case these lead to different values for the same observables, a likely explanation is that new particles contribute to the loop diagrams.

1.3 The neutral B meson system

Two types of neutral B mesons exist, and their flavour eigenstates can be written as:

[3]There is currently a small tension of the 2σ signal level, which is not discussed here.

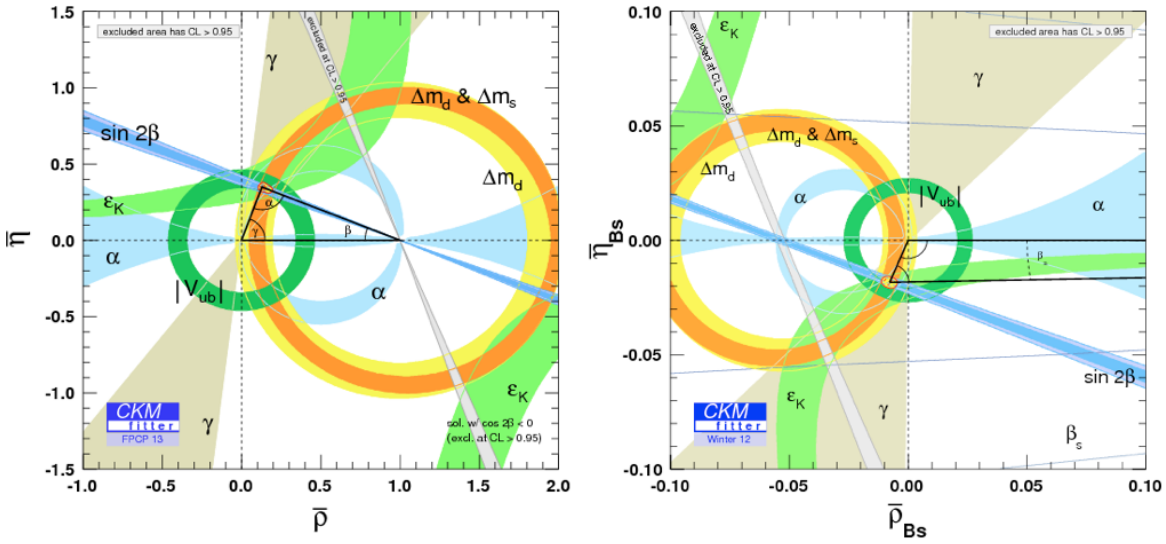

Figure 1.2: *Latest results from the CKMfitter Group for B_d triangle constraints (left) and B_s triangle constraints (right) [27, 83].*

$$|B^0\rangle = |\bar{b}d\rangle \quad , \quad |\bar{B^0}\rangle = |b\bar{d}\rangle$$
$$|B_s^0\rangle = |\bar{b}s\rangle \quad , \quad |\bar{B_s^0}\rangle = |b\bar{s}\rangle$$

Since they contain a d or a $\bar{d}$ quark, the B^0 and $\bar{B^0}$ are referred to as the B_d-system. Similarly, the B_s^0 and $\bar{B_s^0}$ form the B_s-system. The main characteristics of B mesons are, on the one hand, that the mass of these particles is relatively large, due to the fact that a b quark is present in a bound state and, on the other hand, that their lifetime (τ_B) is relatively large due to flavour changing decays. Both the high mass and the long lifetime are signatures used in the detection and identification of B meson decays, since the primary vertex in which the B meson is produced is typically well separated from the secondary (decay) vertex and the decay particles have a transverse momentum larger than those in many background events. In particular, it is important to have a detector able to reconstruct the decay topology at trigger time.

1.3.1 Mixing of neutral B mesons

An important characteristic of neutral B mesons is that they oscillate between the particle and antiparticle state, meaning that a $B_{(s)}^0$ can evolve in time into a $\bar{B}_{(s)}^0$ and vice versa, before decaying. From this point on, the discussion will be presented in terms of B_s^0 mesons[4], as the work of this thesis focuses on these particles. The time evolution of the B_s^0 and $\bar{B}_s^0$ states can be expressed in terms of the flavour eigenstates:

$$|\phi(t)\rangle = a(t)|B_s^0\rangle + b(t)|\bar{B}_s^0\rangle \tag{1.17}$$

The time evolution of the B_s system is described by the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} |a(t)\rangle \\ |b(t)\rangle \end{pmatrix} = H\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{1.18}$$

[4]The same calculations are valid also for B^0 mesons.

where $\mathbf{M}$ and $\mathbf{\Gamma}$ are 2 x 2 hermitian matrices:

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad , \quad \mathbf{\Gamma} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \tag{1.19}$$

The off-diagonal elements of the two matrices, M_{12} and Γ_{12}, are dominated by short-distance (virtual) processes and long distance processes (real intermediate states), respectively. The real intermediate states dominating Γ_{12} are decay states common to both B_s^0 and $\bar{B}_s^0$ mesons. The element M_{12} is dominated by the box diagrams of Fig. 1.3, and could be subject to NP effects. If $M_{12} = \Gamma_{12} = 0$, the B_s^0-meson created in a specific flavour eigenstate would remain in that state indefinitely and no mixing would occur.

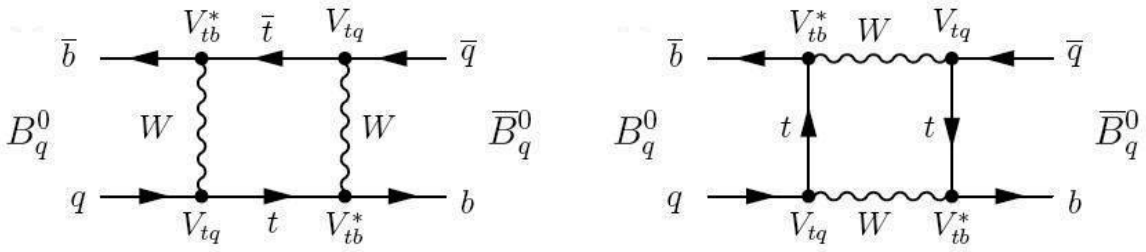

Figure 1.3: *Feynman box diagrams for B_q^0 mixing. The index q indicates either a d or an s quark.*

Since quark flavour is not conserved in weak interactions, M_{12} and Γ_{12} are non-zero and transitions between neutral mesons and their antiparticles are possible. In the B_s-system, this results in the fact that the mass eigenstates of B_s^0 mesons are different from their flavour eigenstates. In fact, the *light* and *heavy* mass eigenstates $|B_L\rangle$ and $|B_H\rangle$, respectively, can be written as a linear combination of the flavour eigenstates as:

$$\begin{aligned} |B_L\rangle &= p|B_s^0\rangle + q|\bar{B}_s^0\rangle \\ |B_H\rangle &= p|B_s^0\rangle - q|\bar{B}_s^0\rangle \end{aligned} \tag{1.20}$$

where $|p|^2 + |q|^2 = 1$. Since $(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma})$ is diagonal in the mass eigenstate basis, the time dependence of the *light* and *heavy* mass eigenstates can then be expressed as:

$$\begin{aligned} |B_L(t)\rangle &= e^{-(im_L + \frac{\Gamma_L}{2})t}|B_L(0)\rangle \\ |B_H(t)\rangle &= e^{-(im_H + \frac{\Gamma_H}{2})t}|B_H(0)\rangle \end{aligned} \tag{1.21}$$

where the masses and widths of the two mass eigenstates are indicated as $m_{L,H}$ and $\Gamma_{L,H}$, respectively.

The difference in mass and lifetime of B_L and B_H for the B_s^0 is written as:

$$\begin{aligned} \Delta m_s &= m_H - m_L, \\ \Delta\Gamma_s &= \Gamma_L - \Gamma_H \end{aligned}$$

In these relations Δm_s is positive by definition and $\Delta\Gamma_s$ can have either sign, but is positive in the SM.
The average lifetime $\tau(B_s^0)$ of the B_s^0 meson is defined according to:

$$\Gamma_s = \frac{1}{\tau(B_s)} = \frac{\Gamma_L + \Gamma_H}{2}$$

By solving the Schrödinger equation 1.18, the ratio q/p is found as a function of the off-diagonal elements M_{12} and Γ_{12}:

$$\frac{q}{p} = -\sqrt{\frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2}} \tag{1.22}$$

The time evolution of the flavour eigenstate $|B_s^0(t)\rangle$ can be expressed in terms of the *light* and *heavy* time-dependent mass eigenstates, as:

$$|B_s^0(t)\rangle = \frac{|B_L(t)\rangle + |B_H(t)\rangle}{2p} =$$

$$= \frac{e^{-(im_L + \frac{\Gamma_L}{2})t} \cdot |B_L(0)\rangle + e^{-(im_H + \frac{\Gamma_H}{2})t} \cdot |B_H(0\rangle)}{2p}$$

and similarly:

$$|\bar{B}_s^0(t)\rangle = \frac{|B_L(t)\rangle - |B_H(t)\rangle}{2q} =$$

$$= \frac{e^{-(im_H + \frac{\Gamma_H}{2})t} \cdot |B_H(0\rangle) - e^{-(im_L + \frac{\Gamma_L}{2})t} \cdot |B_L(0)\rangle}{2q}$$

We define the term:

$$g_\pm(t) = \frac{1}{2}(e^{-(im_L + \frac{\Gamma_L}{2})t} \pm e^{-(im_H + \frac{\Gamma_H}{2})t}),$$

which can also be written as:

$$\begin{aligned} g_+(t) &= e^{-imt}e^{-\Gamma_s t} \cdot \cos\left(\frac{\Delta m_s t}{2}\right), \\ g_-(t) &= e^{-imt}e^{-\Gamma_s t} \cdot i \sin\left(\frac{\Delta m_s t}{2}\right) \end{aligned} \tag{1.23}$$

where it is interesting to note the factor i that gives a 90° phase shift between the $g_+(t)$ and $g_-(t)$ amplitudes.
For a particle that is produced as a B_s^0 meson at $t = 0$ the time evolution of the state is:

$$|B_s^0(t)\rangle = g_+(t)|B_s^0\rangle + \left(\frac{q}{p}\right) g_-(t)|\bar{B}_s^0\rangle, \tag{1.24}$$

while that for a particle that is produced as a $\bar{B}_s^0$ meson at $t = 0$ is:

$$|\bar{B}_s^0(t)\rangle = g_+(t)|\bar{B}_s^0\rangle + \left(\frac{p}{q}\right) g_-(t)|B_s^0\rangle \tag{1.25}$$

Finally, the probability of observing a B_s^0 or $\bar{B}_s^0$ at a time t in case the original particle was produced as either a B_s^0 or $\bar{B}_s^0$ is given by squaring the decay amplitudes:

$$\begin{aligned} &\text{(i)} \quad |\langle B_s^0|B_s^0(t)\rangle|^2 = |g_+(t)|^2 \\ &\text{(ii)} \quad |\langle B_s^0|\bar{B}_s^0(t)\rangle|^2 = \left|\frac{p}{q}\right|^2 |g_-(t)|^2 \\ &\text{(iii)} \quad |\langle \bar{B}_s^0|B_s^0(t)\rangle|^2 = \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \\ &\text{(iv)} \quad |\langle \bar{B}_s^0|\bar{B}_s^0(t)\rangle|^2 = |g_+(t)|^2 \end{aligned} \tag{1.26}$$

with:

$$|g_{\pm}(t)|^2 = \frac{e^{-\Gamma_s t}}{2}\left(\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) \pm \cos\left(\Delta m_s t\right)\right)$$

1.3.2 Decay of neutral B_s mesons

Four decay amplitudes can be defined for a B_s^0 or $\bar{B}_s^0$ that decays directly into a final state f or into the charge-conjugated state $\bar{f}$:

$$\begin{array}{lll} A_f \equiv \langle f|T|B_s^0\rangle & , & \bar{A}_f \equiv \langle f|T|\bar{B}_s^0\rangle \\ A_{\bar{f}} \equiv \langle \bar{f}|T|B_s^0\rangle & , & \bar{A}_{\bar{f}} \equiv \langle \bar{f}|T|\bar{B}_s^0\rangle \end{array} \tag{1.27}$$

where T represents the transition matrix. From the decay amplitudes A_f and $\bar{A}_f$, the parameters λ_f and $\bar{\lambda}_{\bar{f}}$ can be defined as:

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \quad , \quad \bar{\lambda}_{\bar{f}} = \frac{p}{q}\frac{A_{\bar{f}}}{\bar{A}_{\bar{f}}} \tag{1.28}$$

The decay rate of a B_s^0 or $\bar{B}_s^0$ meson which decays into a final state f or $\bar{f}$ follow from Eq. 1.26:

$$\begin{aligned} \Gamma_{B_s^0\to f}(t) = |A_f|^2 \frac{1+|\lambda_f|^2}{2} e^{-\Gamma_s t}. \\ \left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + C_f\cos(\Delta m_s t) - D_f \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - S_f \sin(\Delta m_s t)\right] \end{aligned} \tag{1.29}$$

$$\begin{aligned} \Gamma_{\bar{B}_s^0\to f}(t) = |A_f|^2 \cdot \left|\frac{p}{q}\right|^2 \frac{1+|\lambda_f|^2}{2} e^{-\Gamma_s t}. \\ \left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - C_f\cos(\Delta m_s t) - D_f \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + S_f \sin(\Delta m_s t)\right] \end{aligned} \tag{1.30}$$

$$\begin{aligned} \Gamma_{\bar{B}_s^0\to \bar{f}}(t) = \left|\bar{A}_{\bar{f}}\right|^2 \frac{1+\left|\bar{\lambda}_{\bar{f}}\right|^2}{2} e^{-\Gamma_s t}. \\ \left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + C_{\bar{f}}\cos(\Delta m_s t) - D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) - S_{\bar{f}} \sin(\Delta m_s t)\right] \end{aligned} \tag{1.31}$$

$$\begin{aligned} \Gamma_{B_s^0\to \bar{f}}(t) = \left|\bar{A}_{\bar{f}}\right|^2 \cdot \left|\frac{q}{p}\right|^2 \frac{1+\left|\bar{\lambda}_{\bar{f}}\right|^2}{2} e^{-\Gamma_s t}. \\ \left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - C_{\bar{f}}\cos(\Delta m_s t) - D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + S_{\bar{f}} \sin(\Delta m_s t)\right] \end{aligned} \tag{1.32}$$

where C_f, $C_{\bar{f}}$, D_f, $D_{\bar{f}}$, S_f, $S_{\bar{f}}$ are known as CP coefficients and are given by:

$$\begin{array}{lll} C_f = \dfrac{1-|\lambda_f|^2}{1+|\lambda_f|^2} & , & C_{\bar{f}} = \dfrac{1-\left|\bar{\lambda}_{\bar{f}}\right|^2}{1+\left|\bar{\lambda}_{\bar{f}}\right|^2}, \\ D_f = \dfrac{2\Re\left(\lambda_f\right)}{1+|\lambda_f|^2} & , & D_{\bar{f}} = \dfrac{2\Re\left(\bar{\lambda}_{\bar{f}}\right)}{1+\left|\bar{\lambda}_{\bar{f}}\right|^2}, \\ S_f = \dfrac{2\Im\left(\lambda_f\right)}{1+|\lambda_f|^2} & , & S_{\bar{f}} = \dfrac{2\Im\left(\bar{\lambda}_{\bar{f}}\right)}{1+\left|\bar{\lambda}_{\bar{f}}\right|^2} \end{array} \tag{1.33}$$

By measuring the CP observables C_f, $C_{\bar{f}}$, D_f, $D_{\bar{f}}$, S_f, $S_{\bar{f}}$ of a given decay, the full expression of the decay rate equation is known and a measurement of CP violation can be extracted [28].

1.4 CP Violation in the B_s^0 meson system

Three different types of CP violation can occur in the B_s^0 meson system: *direct* CP violation (also known as CP violation in decay), CP violation in *mixing* and CP violation in the *interference* between decay and mixing.

The possibility of CP violation is included in the SM via complex coupling constants, but a CP violating phenomenon can only occur when at least two amplitudes contribute to the same final state. Moreover, a relative phase difference must exist between these two decay amplitudes A_1 and A_2. This phase difference is made up of two parts: a *strong phase* δ, which does not change sign under CP conjugation, and a *weak phase* ϕ, which does change sign under CP conjugation. By choosing the phase of A_1 to be zero, the ratio between the decay amplitudes A_2 and A_1 can be written in terms of these strong and weak phase differences as:

$$\frac{A_2}{A_1} = \frac{|A_2|}{|A_1|} e^{i\delta} e^{i\phi} \tag{1.34}$$

The overall decay rate can then be written as:

$$\begin{aligned} |A|^2 = |A_1 + A_2|^2 = \left| |A_1| + |A_2| \, e^{i\delta} e^{i\phi} \right|^2 = \\ = |A_1|^2 + |A_2|^2 + 2\,|A_1|\,|A_2| \cos(\delta + \phi) \end{aligned} \tag{1.35}$$

while the CP-conjugated decay rate will be:

$$\begin{aligned} |\bar{A}|^2 = |\bar{A}_1 + \bar{A}_2|^2 = \left| |A_1| + |A_2| \, e^{i\delta} e^{-i\phi} \right|^2 = \\ = |A_1|^2 + |A_2|^2 + 2|A_1||A_2| \cos(\delta - \phi) \end{aligned} \tag{1.36}$$

The CP asymmetry A_{CP} is defined as:

$$A_{CP} = \frac{|A|^2 - |\bar{A}|^2}{|A|^2 + |\bar{A}|^2} = \frac{-2\sin(\delta)\sin(\phi)}{\frac{|A_1|}{|A_2|} + \frac{|A_2|}{|A_1|} + 2\cos(\delta)\cos(\phi)} \tag{1.37}$$

This shows that at least two decay amplitudes must contribute to the final state and that both the relative strong phase δ and the relative weak phase ϕ must be different from zero in order for CP violation to be observed.

1.4.1 Direct CP violation

Direct CP violation takes place when the decay rate of a B_s^0 meson to a final state f is different from the decay rate of a $\bar{B}_s^0$ meson to the CP-conjugated final state $\bar{f}$. In terms of decay amplitudes, this means that:

$$\left| \langle f|T|B_s^0 \rangle \right| \neq \left| \langle \bar{f}|T|\bar{B}_s^0 \rangle \right| \ , \ i.e. \ |A_f| \neq \left| \bar{A}_{\bar{f}} \right| \tag{1.38}$$

This can only occur if the decay amplitude A_f contains two interfering amplitudes (A_1 and A_2 as before), where e.g. the first amplitude can be a tree diagram and the second a so-called penguin (loop) diagram to which perhaps virtual new particles can contribute.

As can be deduced from Eq. 1.37, the largest CP violation occurs when the two amplitudes are of equal magnitude and have a weak and strong phase difference of 90°. In practice, the weak phase ϕ is the phase difference between CKM elements in the two decay amplitudes and the strong phase δ originates from strong interactions in the final state.

1.4.2 CP violation in mixing

CP violation in mixing takes place if the probability of a B_s^0 meson to oscillate to a $\bar{B}_s^0$ meson is different from the probability of a $\bar{B}_s^0$ meson to oscillate to a B_s^0 meson. Using (*ii*) and (*iii*) of the relations defined in Eq. 1.26, this can be written as:

$$\left|\langle B_s^0|\bar{B}_s^0(t)\rangle\right|^2 = \left|\frac{p}{q}\right|^2 |g_-(t)|^2 \neq \left|\frac{q}{p}\right|^2 |g_-(t)|^2 = \left|\langle \bar{B}_s^0|B_s^0(t)\rangle\right|^2 \tag{1.39}$$

Therefore, CP violation in mixing is present when $|p/q| \neq |q/p|$ or, equivalently, when:

$$\left|\frac{q}{p}\right| \neq 1 \tag{1.40}$$

From Eq. 1.22 it can be seen that, in order for this condition to be satisfied, it must be that $|M_{12}^* - i\Gamma_{12}^*/2| \neq |M_{12} - i\Gamma_{12}/2|$ or consequently that $|(M_{12} - i\Gamma_{12}/2)^*| \neq |(M_{12} - i\Gamma_{12}/2)|$, which is satisfied if there is a $non - 0^\circ$ or $non - 180^\circ$ phase difference between M_{12} and Γ_{12}.

In the B^0 and B_s^0 system, $|\Gamma_{12}| << |M_{12}|$ and $|q/p| \simeq 1$ both within experimental accuracy. Instead, CP violation in mixing has been observed in the kaon system.

1.4.3 CP violation in the interference between decays with and without mixing

The third type of CP violation that can occur is CP violation in the interference between decays with and without mixing. In order for this type of CP violation to be present, a final state f must exist to which a B_s^0 meson can either decay directly or after having first oscillated into a $\bar{B}_s^0$ meson; the amplitudes are $A(B_s^0 \to f)$ and $A(B_s^0 \to \bar{B}_s^0 \to f)$, respectively. If there is a relative weak phase difference in the mixing amplitude or between the amplitude of the direct decay and the amplitude of the decay after mixing has occurred, then CP violation in interference can be observed.

The strong phase is provided by the mixing mechanism which includes a 90° phase shift between $g_+(t)$ and $g_-(t)$, as was made explicit in Eq. 1.23.

1.5 The decays $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$

In this thesis, the B_s^0 meson decays $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ are investigated and a method with which the CKM angle γ can be extracted is described.

1.5.1 The decay $B_s^0 \to D_s^- \pi^+$

The decay $B_s^0 \to D_s^- \pi^+$ occurs through only one tree diagram, as shown in Fig. 1.4, while the decay $\bar{B}_s^0 \to D_s^- \pi^+$ is not allowed.

Decays of an initially produced B_s^0 or $\bar{B}_s^0$ to the charge-conjugated final states can only occur after mixing. This means that for these decays $A_{\bar{f}} = \bar{A}_f = 0$, which in turn implies $\lambda_f = \bar{\lambda}_{\bar{f}} = 0$ for the decay equations 1.29 - 1.32. If we further assume that there is no CP violation in mixing in the B_s^0 system (i.e. $|q/p| = 1$), then the CP coefficients of Eq. 1.33 are given by:

$$C_f = C_{\bar{f}} = 1 \ ,$$

$$D_f = S_f = D_{\bar{f}} = S_{\bar{f}} = 0 \ ,$$

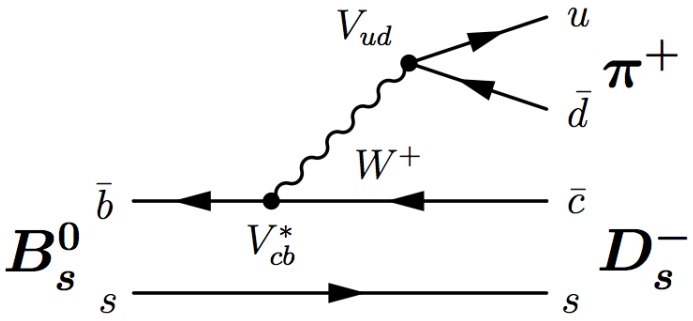

Figure 1.4: *Feynman diagram of the decay* $B_s^0 \to D_s^- \pi^+$.

1.5.2 The decay $B_s^0 \to D_s^\mp K^\pm$

The decay $B_s^0 \to D_s^\mp K^\pm$ can occur directly via a tree diagram but also via mixing and an alternative tree diagram, indicated in Fig. 1.5.

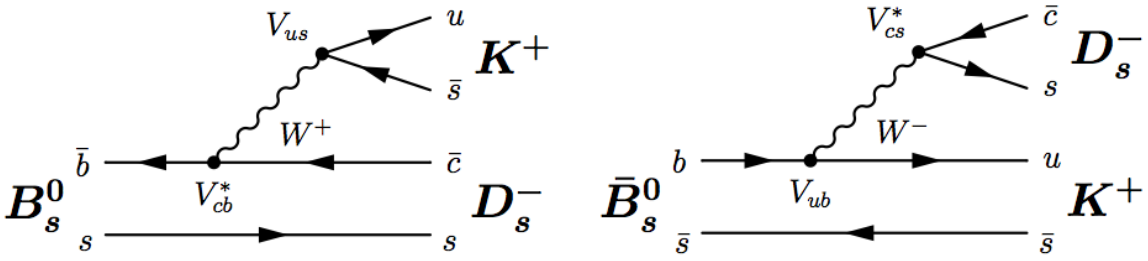

Figure 1.5: *Feynman diagrams of the decays* $B_s^0 \to D_s^- K^+$ *and* $\bar{B}_s^0 \to D_s^- K^+$.

The presence of a direct tree diagram and a tree decay diagram with mixing means that the two amplitudes contributing to this decay interfere, implying that CP violation might be observed in this channel (see Sec. 1.4).

Assuming there is no direct CP violation in the decay, $|A_f| = |\bar{A}_{\bar{f}}|$ and $|A_{\bar{f}}| = |\bar{A}_f|$, which together with $|q/p| = 1$ leads to $|\lambda_f| = |\bar{\lambda}_{\bar{f}}|$. For the decay $B_s^0 \to D_s^\mp K^\pm$, the CP coefficients can be expressed as:

$$C_f = C_{\bar{f}} = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}$$

$$D_f = \frac{2|\lambda_f| \cdot \cos\phi_\lambda}{1 + |\lambda_f|^2} \quad , \quad \mathrm{D}_{\bar{f}} = \frac{2|\bar{\lambda}_{\bar{f}}| \cdot \cos\phi_{\bar{\lambda}}}{1 + |\bar{\lambda}_{\bar{f}}|^2} \tag{1.41}$$

$$S_f = \frac{-2|\lambda_f| \cdot \sin\phi_\lambda}{1 + |\lambda_f|^2} \quad , \quad \mathrm{S}_{\bar{f}} = \frac{-2|\bar{\lambda}_{\bar{f}}| \cdot \sin\phi_{\bar{\lambda}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}$$

The CKM angle γ can be measured in the decay $B_s^0 \to D_s^\mp K^\pm$ [**?**]. When the flavour of the initial B_s^0 meson is known, all of the CP observables can be extracted from the decay rate equations, allowing for a measurement of CP violation. In case the flavour of the B_s^0 meson in

its initial state is not known the decay rate can be written as:

$$\begin{aligned}\Gamma_{B_s\to f}(t) + \Gamma_{\bar{B}_s\to f}(t) = \\ = \left(1 + \left|\frac{p}{q}\right|^2\right)|A_f|^2(1+|\lambda_f|^2)\; e^{-\Gamma_s t}\;\left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - D_f \sinh\left(\frac{\Delta\Gamma t}{2}\right)\right]\end{aligned} \tag{1.42}$$

and a measurement of CP violation can still in principle be obtained through determination of the CP observables D_f and $\bar{D}_{\bar{f}}$.

From Fig. 1.5 it can be seen that the amplitude of the decay $B_s^0 \to D_s^- K^+$ (on the left in the figure) is proportional to $V_{cb}^* V_{us}$, while the amplitude of the decay $B_s^0 \to \bar{B}_s^0 \to D_s^- K^+$ (on the right) is proportional to $V_{ub} V_{cs}^*$.

Factorizing out the CKM elements, the ratio between $\bar{A}_{D_s^- K^+}$ and $A_{D_s^- K^+}$ can be written as:

$$r_{D_s K} = \left|\frac{\bar{A}_{D_s^- K^+}}{A_{D_s^- K^+}}\right| = \left|\frac{V_{ub}V_{cs}^*}{V_{cb}^* V_{us}} \cdot \frac{A_2}{A_1}\right| \simeq 0.37 \tag{1.43}$$

where A_2/A_1 is the ratio of the *hadronic* amplitudes that contribute to the final state and which are expected to be close to unity. The ratio between the decay amplitudes $\bar{A}_{D_s^- K^+}$ and $A_{D_s^- K^+}$ can be written as:

$$\left(\frac{\bar{A}_{D_s^- K^+}}{A_{D_s^- K^+}}\right) = \frac{|\bar{A}_{D_s^- K^+}|}{|A_{D_s^- K^+}|} e^{i(\delta-\gamma)} \tag{1.44}$$

where $\gamma = \arg\left(V_{ub}V_{cs}^*/V_{cb}^*V_{us}\right)$ and δ the relative phase of the hadronic amplitudes.

The parameter λ_f follows by factorizing in the amplitude of the box diagram of Fig. 1.3.

$$\lambda_{D_s^- K^+} = \left(\frac{q}{p}\right)_{B_s^0}\left(\frac{\bar{A}_{D_s^- K^+}}{A_{D_s^- K^+}}\right) = \left|\frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*}\right|\left|\frac{V_{ub}V_{cs}^*}{V_{cb}^* V_{us}}\right|\left|\frac{A_2}{A_1}\right| e^{i(-2\beta_s-\gamma+\delta)} \tag{1.45}$$

where $2\beta_s$ is the additional relative phase of the mixing diagram.

The combined CKM angle $2\beta_s + \gamma$ can then be extracted by combining the expression of $\lambda_{D_s^- K^+}$ for the pair of decays $B_s^0 \to D_s^- K^+$ and $B_s^0 \to \bar{B}_s^0 \to D_s^- K^+$, and the expression of $\lambda_{D_s^+ K^-}$ for the pair of decays to the CP-conjugated final state $\bar{B}_s^0 \to D_s^+ K^-$ and $\bar{B}_s^0 \to B_s^0 \to D_s^+ K^-$. For these decays $\bar{\lambda}_{D_s^+ K^-}$ is obtained in a way similar to the case of $\lambda_{D_s^- K^+}$, and is given by:

$$\bar{\lambda}_{D_s^+ K^-} = \left(\frac{p}{q}\right)_{B_s^0}\left(\frac{A_{D_s^+ K^-}}{\bar{A}_{D_s^+ K^-}}\right) = \left|\frac{V_{tb}V_{ts}^*}{V_{tb}^* V_{ts}}\right|\left|\frac{V_{us}V_{cb}^*}{V_{cs}^* V_{ub}}\right|\left|\frac{A_1}{A_2}\right| e^{i(2\beta_s+\gamma+\delta)} \tag{1.46}$$

The result is a determination of the sum $(\delta + (2\beta_s + \gamma))$ and difference $(\delta - (2\beta_s + \gamma))$ of the weak and strong phases, and therefore of each individually.

In terms of the weak and strong phases δ and $(2\beta_s + \gamma)$, the CP coefficients are written explicitely as:

$$C_f = C_{\bar{f}} = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}$$

$$D_f = \frac{2|\lambda_f|\cdot\cos(\delta-(2\beta_s+\gamma))}{1+|\lambda_f|^2} \quad , \quad \mathrm{D}_{\bar{f}} = \frac{2|\bar{\lambda}_{\bar{f}}|\cdot\cos(\delta+(2\beta_s+\gamma))}{1+|\bar{\lambda}_{\bar{f}}|^2} \tag{1.47}$$

$$S_f = \frac{-2|\lambda_f|\cdot\sin(\delta-(2\beta_s+\gamma))}{1+|\lambda_f|^2} \quad , \quad \mathrm{S}_{\bar{f}} = \frac{-2|\bar{\lambda}_{\bar{f}}|\cdot\sin(\delta+(2\beta_s+\gamma))}{1+|\bar{\lambda}_{\bar{f}}|^2}$$

These are the parameters that are fitted in the decay rates in Ch. 6.

1.6 Experimental effects that affect the measurement of B_s^0 oscillations

The measurement of the angle γ involves the determination of the time-dependent decay rate constants C_f, D_f, $D_{\bar{f}}$, S_f and $S_{\bar{f}}$, for the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay rate. The observables are correlated to the amplitude of the decay rate oscillations and can potentially be affected by experimental effects. These effects are considered in this section.

B_s^0 oscillations occur at a high frequency of $17.8 \cdot 10^{12}$ Hz[5], which corresponds to a distance per oscillation of about 100 μm in the LHC. These oscillations can be observed by LHCb but are diluted by experimental effects such as proper time resolution and an imperfect flavour tagging of the B_s^0 mesons at production. In Fig. 1.6, the proper time distribution for a $B_s^0 \to D_s^- \pi^+$ decay is shown in the case of perfect decay time resolution and perfect tagging.

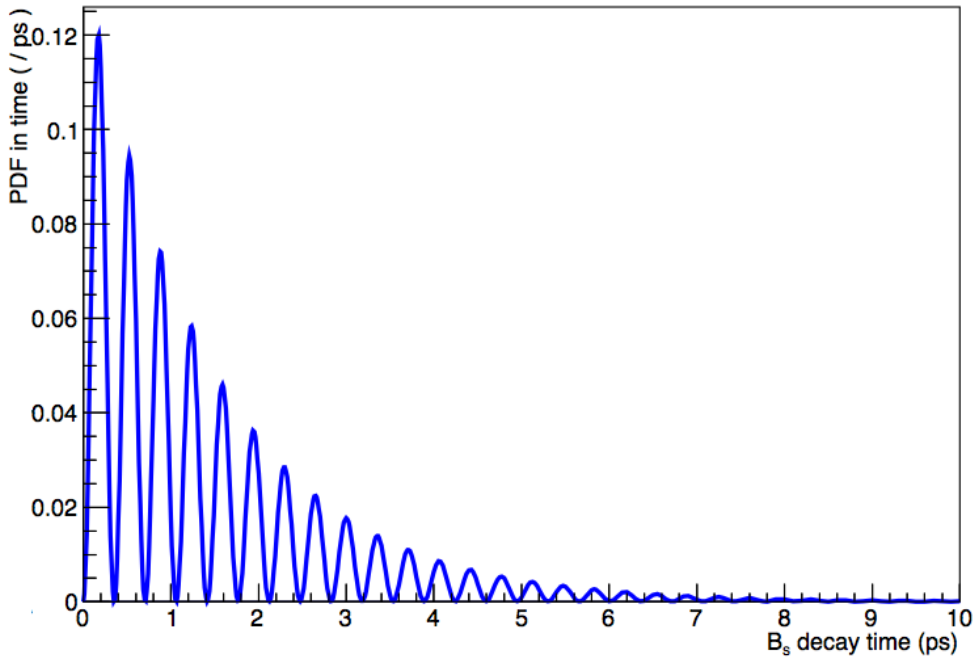

Figure 1.6: *Ideal proper time distribution for a $B_s^0 \to D_s^- \pi^+$ decay, in the case of perfect proper time resolution and perfect tagging efficiency (with Δm_s=17.76 ps^{-1}).*

1.6.1 Realistic proper time resolution performance

The finite decay time resolution of the detector leads to a dilution of the observed oscillation. In Fig. 1.7, the proper time distribution is shown in case of a Gaussian proper time resolution with σ_t=40 fs. The dilution factor can be extracted by calculating the convolution integral of an oscillation of frequency Δm_s:

$$f(t') = e^{(i\Delta m_s t')} \ ,$$

with a Gaussian resolution function:

$$R(t, t'; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma_{t'}^2}} \cdot e^{-\frac{(t-\mu-t')^2}{2\sigma_{t'}^2}}$$

[5]the latest measurement for this frequency is Δm_s=17.76 ps^{-1} [30].

Here, t' is the true lifetime and t is the time smeared by the resolution model $R(t, t'; \mu, \sigma)$. The convolution is then:

$$\int dt' f(t') \cdot R(t, t'; \mu, \sigma)$$

which finally results in the dilution factor:

$$D_{res} = e^{-\left(\frac{1}{2}\Delta m_s^2 \sigma_t^2\right)},$$

a proper time resolution of σ_t=40 fs leads to a dilution factor of $D_{res} \simeq 0.8$. It can be seen in the figure that the amplitude of the oscillation is instead reduced by the expected dilution factor. Since the observable CP violation is proportional to the amplitude of the oscillations, a correct knowledge of the decay time resolution is required to extract the correct CP observables.

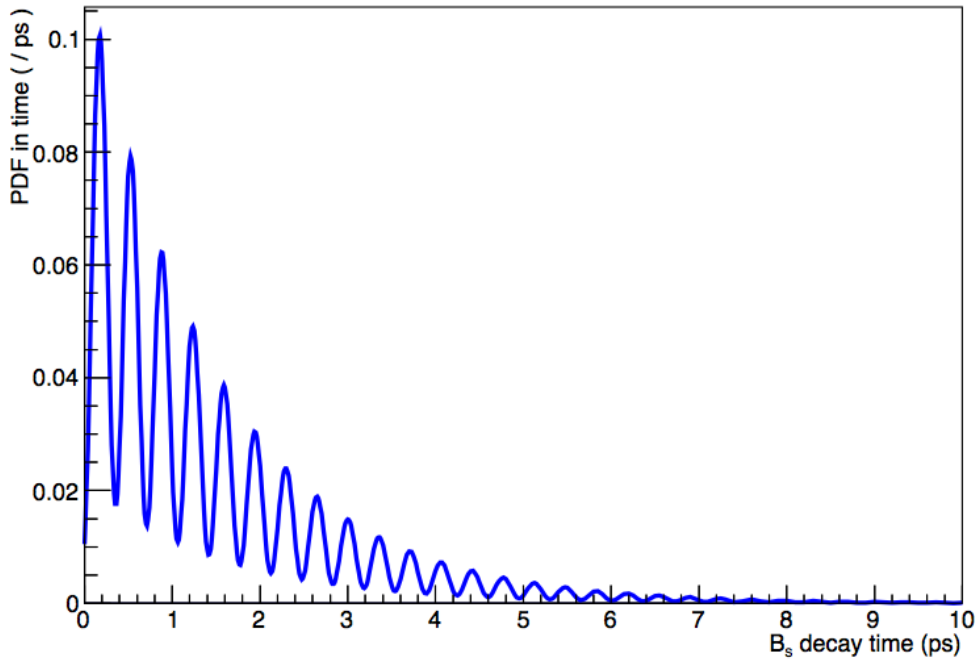

Figure 1.7: *Proper time distribution for a $B_s^0 \to D_s^- \pi^+$ decay, in the case of a finite proper time resolution (σ_t=40 fs).*

1.6.2 Realistic flavour tagging performance

To observe an oscillation signal, the flavour of the B_s^0 at production must be known, since the untagged decay time distribution is the sum of two exponents (see Eq. 1.42). A process called flavour tagging [31] is used to assign tags to the B candidates. However, not always the correct flavour is assigned to the initially produced B_s^0 meson. Tagging algorithms provide the initial flavour of the B_s^0 candidates on an event-per-event basis, as follows:

$$q_i = \begin{cases} 0 \text{ if the candidate is untagged} \\ +1 \text{ if the candidate is tagged as a } B_s^0 \text{ at production} \\ -1 \text{ if the candidate is tagged as a } \bar{B}_s^0 \text{ at production} \end{cases}$$

There are two main classes of flavour tagging: opposite side (OS) tagging, in which the flavour of the non-signal b hadron at production is determined, and same side kaon (SSK) tagging, in which the charge of the fragmentation kaon accompanying the signal B_s^0 is determined.

However, the assigned tag is not necessarily correct. The mistag probability ω gives the fraction of wrong tags among the tagged events. The dilution D_{tag} in terms of the mistag probability is defined as:

$$D_{tag} \;=\; 1 - 2\omega \tag{1.48}$$

The CP coefficients that appear in the time-dependent decay rate equations of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays are measured through their product with the dilution $D = D_{res} \cdot D_{tag}$. To determine the CP coefficients the mistag probability ω as well as the resolution σ_t must be determined precisely.

The effective statistical power of a tagged sample (or tagging power), which indicates the reduction in size of an effective data sample when compared to a perfectly tagged one, is given by:

$$Q_{tag} \;=\; \epsilon_{tag} \cdot (1 - 2\omega)^2 \tag{1.49}$$

In Fig. 1.8, the proper time distribution for a $B_s^0 \to D_s^- \pi^+$ decay is shown in case $\omega \simeq 0.3609$ (the value achieved by LHCb on 2011 data). As shown in the figure, the amplitude of the oscillations is further reduced to about $D = D_{res} \cdot D_{tag} = 0.8 \cdot 0.28 = 0.22$, and a precise determination of the CP parameters becomes more challenging.

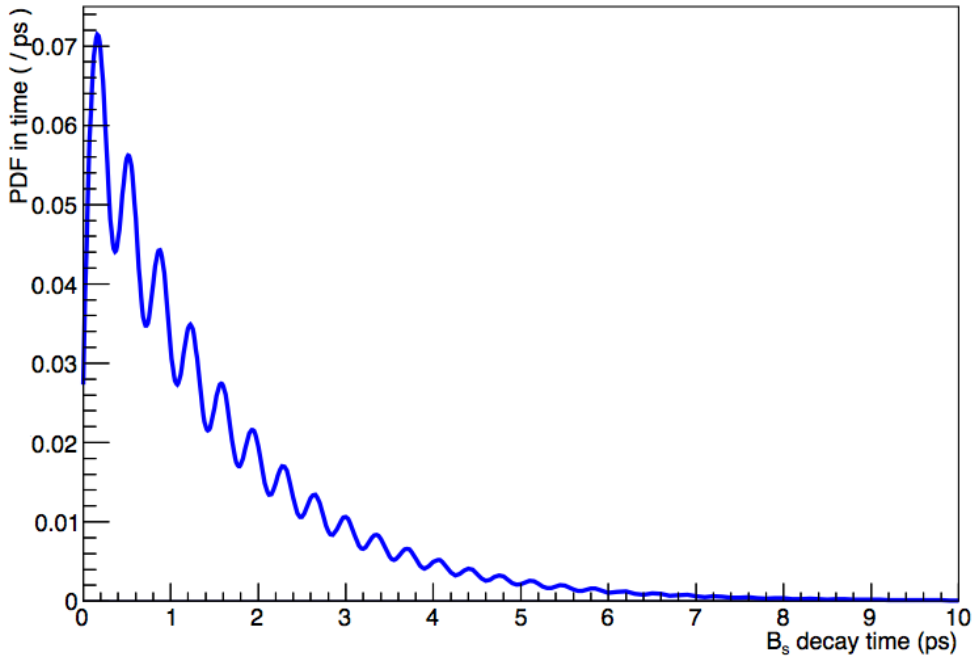

Figure 1.8: *Proper time distribution for a $B_s^0 \to D_s^- \pi^+$ decay, in case ω=0.3609, σ_t=40 fs.*

To gain optimal sensitivity, a neural network (NN) predicts each signal candidate's mistag probability η. Of course, in the optimal case, the predicted value η equals the real value ω for each candidate. However, this optimal case can only be achieved in reality, by calibrating the prediction as the function $\eta(\omega)$. This is the topic of Ch. 7 of this thesis.

Chapter 2

The LHCb Experiment

The main purpose of the LHCb experiment is to study CP-violation in B meson decays as well as rare decays. These studies will lead to a more precise determination of parameters in the quark sector of the Standard Model (SM), and might also give evidence of physics beyond the SM, also known as New Physics (NP). In this chapter, an overview of the LHCb experiment is given. The layout of the Large Hadron Collider (LHC) and the LHCb spectrometer are described in Sec. 2.1 and Sec. 2.2, respectively. In Sec. 2.3 the tracking system of LHCb is described, the particle identification system in Sec. 2.4, and finally the Trigger system in Sec 2.5.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [32] is a proton-proton collider[1] designed to operate at a center of mass energy of $\sqrt{s} = 14$ TeV, a collision frequency of 40 MHz, and a luminosity of 10^{34} cm^{-2}s^{-1}. It is a circular accelerator located in the former LEP tunnel about 100 m underground, with a circumference of approximately 27 km. Along this circumference the four interaction points can be found, at which the four main experiments (ATLAS, CMS, LHCb, ALICE) have been installed. A schematic drawing of the LHC accelerator complex is shown in Fig. 2.1.

The protons pass through several pre-accelerators (LINAC, PSB, PS, SPS) and are injected into the LHC at an energy of 450 GeV. The LHC is a synchrotron having two distinct beam pipes, in which protons are accelerated in opposite directions. Beams 1 and 2 are injected in proximity of the ALICE and LHCb experiments, respectively, as shown in Fig. 2.1 ($TI2$ and $TI8$ in the figure). The curvature in the particle trajectories is obtained using more than 1600 15 m long superconducting magnets, whose coils are kept at a cryogenic temperature of 1.9 K and generate a magnetic field of about 8.4 T. Since the beginning of operations in 2010, the LHC has been delivering collisions with increasing luminosities. During the 2010 and 2011 runs, the proton-proton center of mass energy, $\sqrt{s}$, was equal to 7 TeV, while in 2012 the LHC operated at $\sqrt{s}$ =8 TeV. It is expected to operate at the design energy of $\sqrt{s}$ =14 TeV after the first long shutdown (LS1), which ends early 2015.

The LHC beams are divided in bunches each containing about 10^{11} protons. In 2009 only two proton bunches collided, which steadily increased to 1320 bunches in 2011 and 1380 in 2012. The bunch spacing was decreased from 75 ns in 2010 to 50 ns during the 2011 and 2012 runs. At its design luminosity, the LHC will be filled with 2808 bunches, with a bunch spacing of 25 ns.

[1] It can also operate using heavy ions, leading to Pb-Pb or p-Pb collisions.

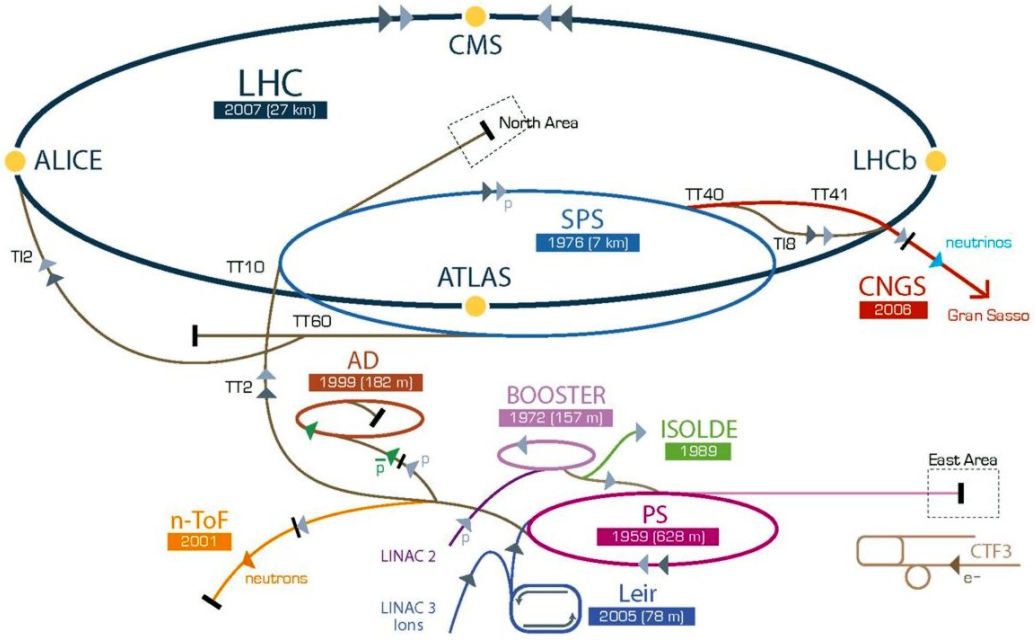

Figure 2.1: *A schematic drawing of the LHC accelerator complex. The four main experiments (ATLAS, CMS, LHCb, ALICE) are located at the four interaction points.*

2.2 The LHCb Experiment

At $\sqrt{s}$= 7 TeV, the total $b\bar{b}$ production cross-section, $\sigma_{b\bar{b}}$, is about 300 μb [33]. This corresponds to about one $b\bar{b}$ pair being produced every 230 pp interactions. In the data taking periods of 2010, 2011 and 2012, LHCb collected 37 pb^{-1}, 1.1 fb^{-1} and 2.1 fb^{-1} of data, respectively. This corresponds to a total of about $1 \cdot 10^{12}$ $b\bar{b}$ pairs produced at the LHCb interaction point. The recorded integrated luminosity for 2010, 2011 and 2012 is shown in Fig. 2.2.
At high energies such as those reached at the LHC, b hadrons are mainly produced in the same forward or backward cone around the beam axis. The fact that the $b\bar{b}$ pairs are typically strongly boosted along one of the two beam directions is due to the fact that the momenta of the partons participating in a pp collision are distributed according to a relatively wide parton distribution function. The LHCb spectrometer is therefore built as a single-arm forward spectrometer. A schematic side view of LHCb is shown in Fig. 2.3.

To collect most of the produced particles in a cone around one beam, LHCb covers a pseudo-rapidity range $2 \leq \eta \leq 5$, where $\eta = -log(tan\frac{\theta}{2})$ and θ is the polar angle with respect to the beam axis. In terms of θ, the LHCb spectrometer has a coverage in the range [10, 300] mrad in the (x, z)-plane, and [10, 250] mrad in the (y, z)-plane and captures about 25% of the produced $b\bar{b}$ pairs. LHCb operates at an optimized luminosity of 4×10^{32} cm^{-2}s^{-1}. The chosen luminosity is lower than the nominal luminosity delivered by the LHC machine in order to limit pile-up: the number of simultaneous pp collisions in the same bunch crossing. Since the interaction rate is too high for storage, further data rejection is needed to increase the fraction of interesting events in the data sample. For this purpose, a two-step trigger is used, made up of a hardware based trigger (Level-0 Trigger) and a software based trigger (High Level Trigger).

In the LHC, B hadrons typically fly about 1 cm before they decay, and a precise determination of the primary and secondary vertex positions is required in order to accurately measure time-dependent decay rates. The vertex detector has a typical transverse impact parameter (IP) resolution parametrized as $11.6 + 23.4/p_T$ μm (see Fig. 2.4). For $B_s \to D_s\pi$ decays for

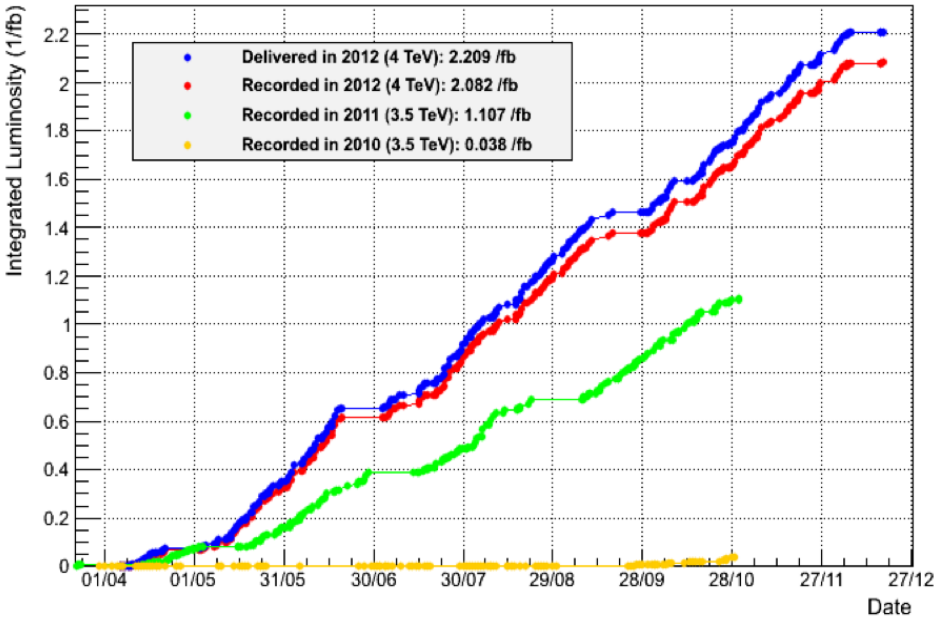

Figure 2.2: *Integrated luminosity during the data-taking periods of 2010, 2011 and 2012.*

example, this results in a typical proper time resolution of 40 fs, which is sufficient to measure time-dependent B_s oscillations (with a typical oscillation period of 350 fs). The complete tracking system together with a 4 Tm dipole magnet yields a momentum resolution of approximately

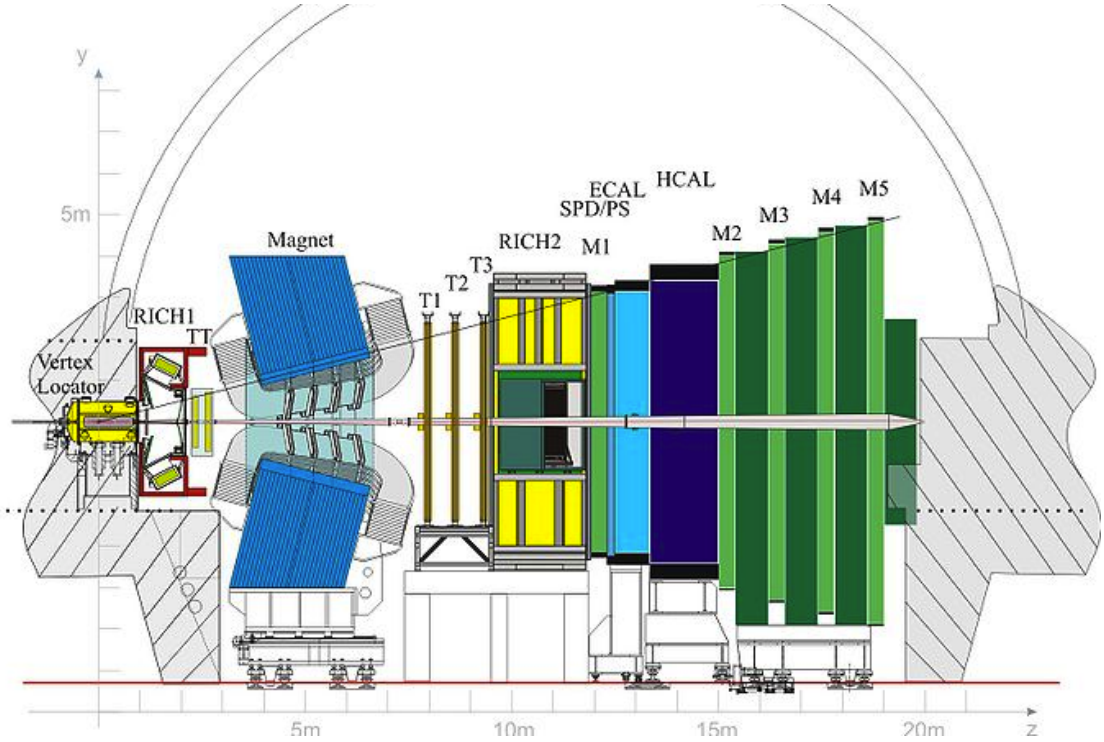

Figure 2.3: *Schematic side view of the LHCb spectrometer.*

$\delta p/p \simeq 0.3 - 0.5\%$ in the momentum range from 5 GeV to 100 GeV, leading to a typical mass resolution of 22 MeV/c^2 for two-body B-decays such as $B \to \pi\pi$ and 14 MeV/c^2 for four-body decays such as $B_s \to D_s\pi$.

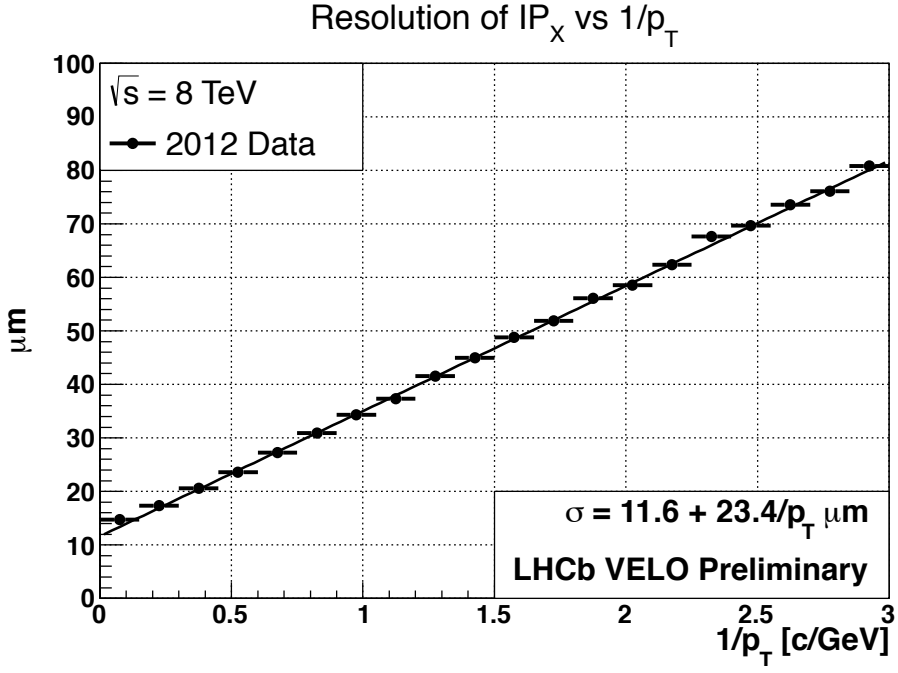

Figure 2.4: *Transverse impact parameter resolution.*

Particle identification is provided by two Ring Imaging CHerenkov (RICH) detectors, which yield π/K separation over a momentum range from 2 GeV to 100 GeV, an electromagnetic (ECAL) and hadronic (HCAL) calorimeter system, which separates photons, electrons and pions, and the muon system consisting of five separate stations, identifying muons.

In the following sections the various components of the LHCb spectrometer are described in more detail.

2.3 Tracking System

The tracking system is located on both sides of the dipole magnet. In front of the magnet, are the Vertex Locator and the Tracker Turicensis, and behind the magnet the Inner Tracker and the Outer Tracker.

2.3.1 Vertex Locator

The VErtex LOcator (VELO) is a silicon strip detector made up of two movable halves used to reconstruct the positions of the primary production and secondary decay vertices [34]. In Fig. 2.5, one half of the detector is shown including the base, the sensors and the RF encapsulation.
Each VELO half has 21 modules with two 300 μm thick sensors mounted back-to-back, each in turn consisting of 2048 silicon strips, one with r-strips to measure the radial r-coordinate and one with ϕ-strips to measure the azimuthal ϕ-coordinate. In addition, two dedicated r-modules at the upstream detector end are used in the Level 0 Pile-up veto detector. The average occupancy of a VELO sensor is below 1%. The R-sensors consist of semicircular silicon strips segmented into four 45° sectors, each with 512 strips. The strip pitch increases linearly as a function of

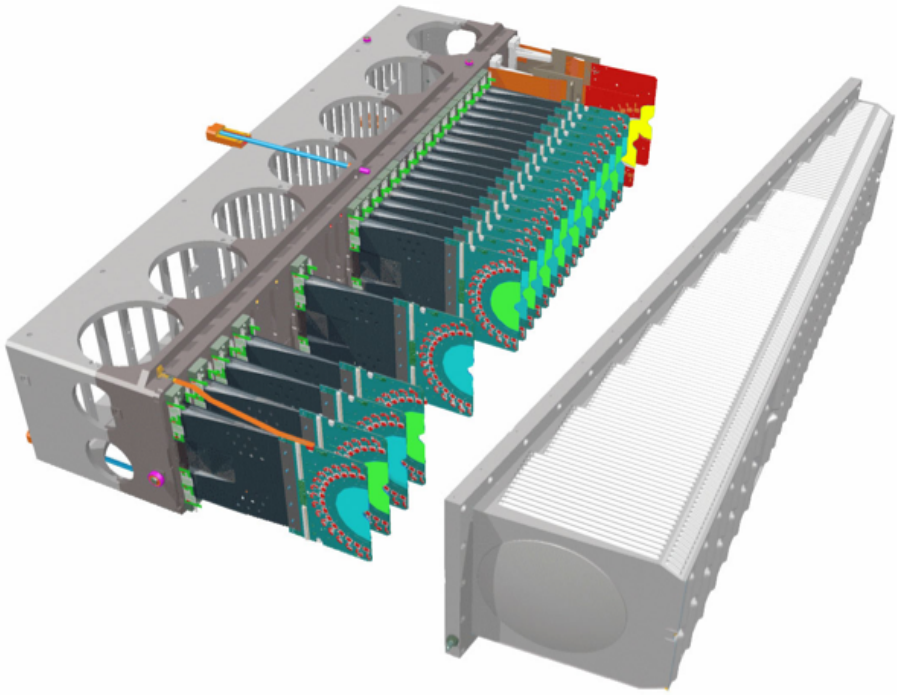

Figure 2.5: *A schematic view of the LHCb VELO showing the base, sensors and RF encapsulation for one half of the detector.*

the radius, corresponding to 38 μm at the inner edge of the sensor and about 102 μm at the outer edge. The ϕ-sensors instead consist of straight silicon strips, divided into an inner and an outer region in which the strips are skewed in opposite directions. The outer region has twice as many strips as the inner region. Also for the ϕ-sensors the strip pitch increases linearly as a function of the radius, ranging from 38 μm$-$78 μm in the inner region and 37 μm$-$97 μm in the outer region. The inner strips are placed at a 20° angle with respect to the radial, while the outer strips are at an angle of 10°. The ϕ-sensors are mounted such that adjacent sensors face opposite directions, such that the strips in adjacent sensors have a stereo angle with respect to each other. This measure improves the pattern recognition capabilities of the detector. A schematic picture of an $R-$ and a ϕ-sensor is shown in Fig. 2.6.

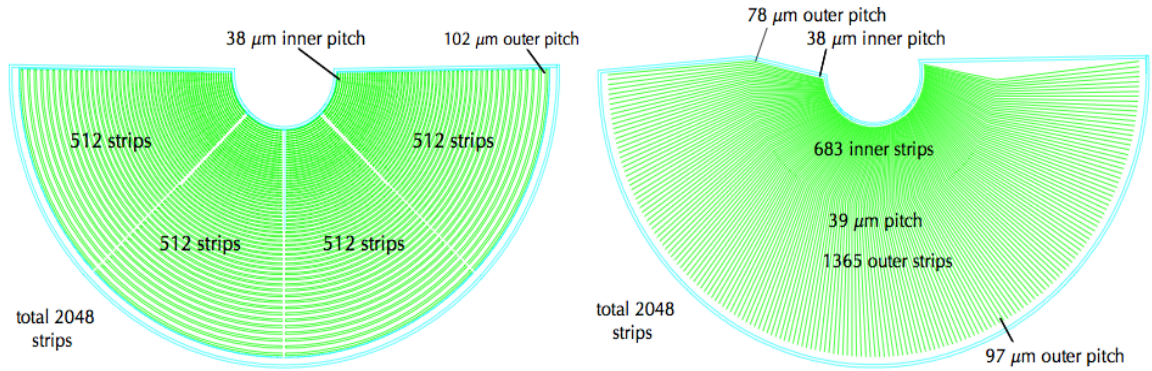

Figure 2.6: *Schematic view of the two sensors of a VELO module, one with R (left) and one with ϕ (right) geometry.*

The vertex resolution provided by the VELO provides a good lifetime resolution for B hadron lifetime measurements. Primary vertices (PV) are found by requiring that a high multiplicity of reconstructed tracks originates from a common point on the LHC beam axis. The PV resolution is determined by splitting a track sample randomly in two and comparing the reconstructed vertex of each set of tracks. Secondary vertices are found using the mass and decay direction requirements of specific B-candidates. The differences between the individually reconstructed vertices are then plotted in bins of the vertex-multiplicity, resulting in one distribution for each track multiplicity. The PV resolution for a specific track multiplicity is then given by the width of the corresponding distribution. As expected, the obtained PV resolution strongly depends on the track multiplicity, as shown in Fig. 2.7. The curves are fitted using the parametrization $(A/N^B)+C$, where N is the track multiplicity. The x and y resolutions are found to be similar. The fitted parameters A, B and C are given in the plots.

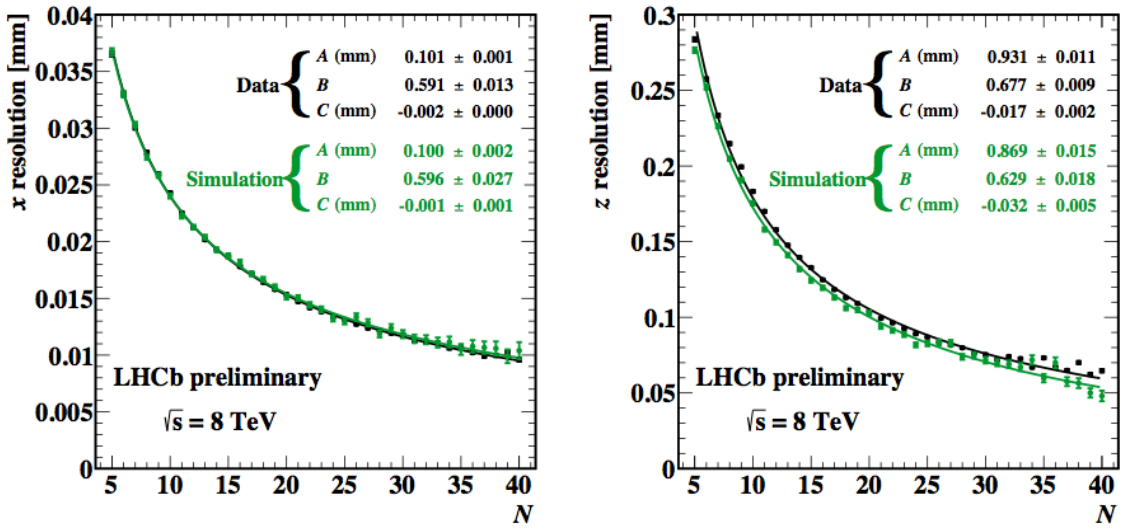

Figure 2.7: *Primary vertex resolution as a function of the number of tracks, in the transverse x-direction (left) and longitudinal z-direction (right). Measurements using 2012 MC and real data are shown in green in black, respectively.*

For a typical vertex of 25 tracks, the PV resolution is equal to about 13 μm in both the x and y-direction and to 85 μm in the z-direction. A more detailed description of the vertex detector, with emphasis on the data acquisition system can be found in Ch. 3.

2.3.2 Tracker Turicensis

The Tracker Turicensis (TT) is a silicon tracker located just upstream of the magnet; it consists of two stations with two layers of silicon strip sensors each. Each sensor covers an area of 7.74 cm $\times$ 7.4 cm and contains 512 Si strips, with a pitch of 183 μm. The four layers are arranged in a $xuvx$ configuration in order to allow for a 3-dimensional coordinate measurement. The first and last layers have their measurement direction along the x-axis, the second is rotated over an angle of -5° around the z-axis, and the third is rotated by $+5^\circ$. The layout of the first two TT layers is shown in Fig. 2.8.

The TT detector serves various purposes in the experiment. It is used to reconstruct tracks that leave no (or insufficient) hits in the VELO, called downstream tracks, as for example those belonging to K_s and Λ decay particles. It is also used to reconstruct low momentum particles, such as slow pions, that do not reach the downstream tracker. In addition, it is used in the online (HLT) reconstruction as a seed for searching for long tracks. Finally, the TT detector is

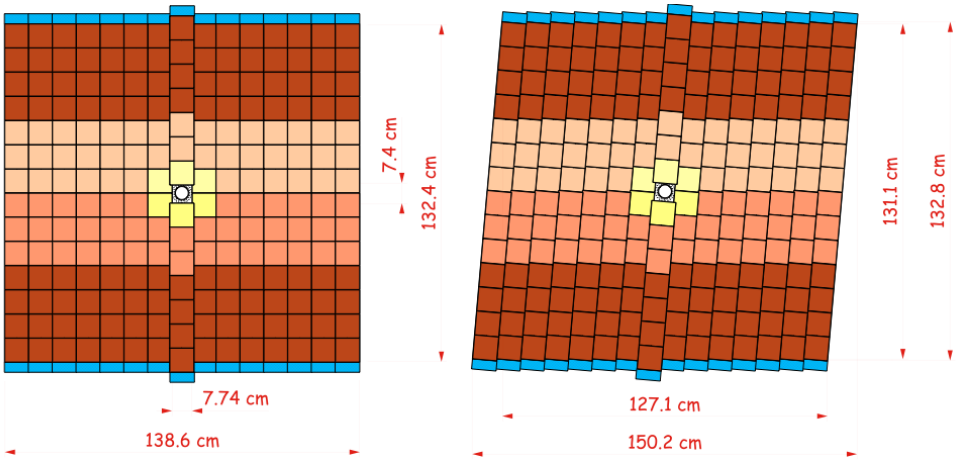

Figure 2.8: *Layout of the first two layers of the Tracker Turicensis.*

used to reject ghost-track combinations of VELO track seeds and downstream track seeds, and to improve the momentum resolution for long tracks.

2.3.3 Dipole Magnet

To allow the measurement of the momentum of charged particles, the LHCb uses a warm dipole magnet that produces an integrated magnetic field of approximately 4 Tm, with its main component along the y-axis. In Fig. 2.9 a photograph of the dipole magnet is shown.

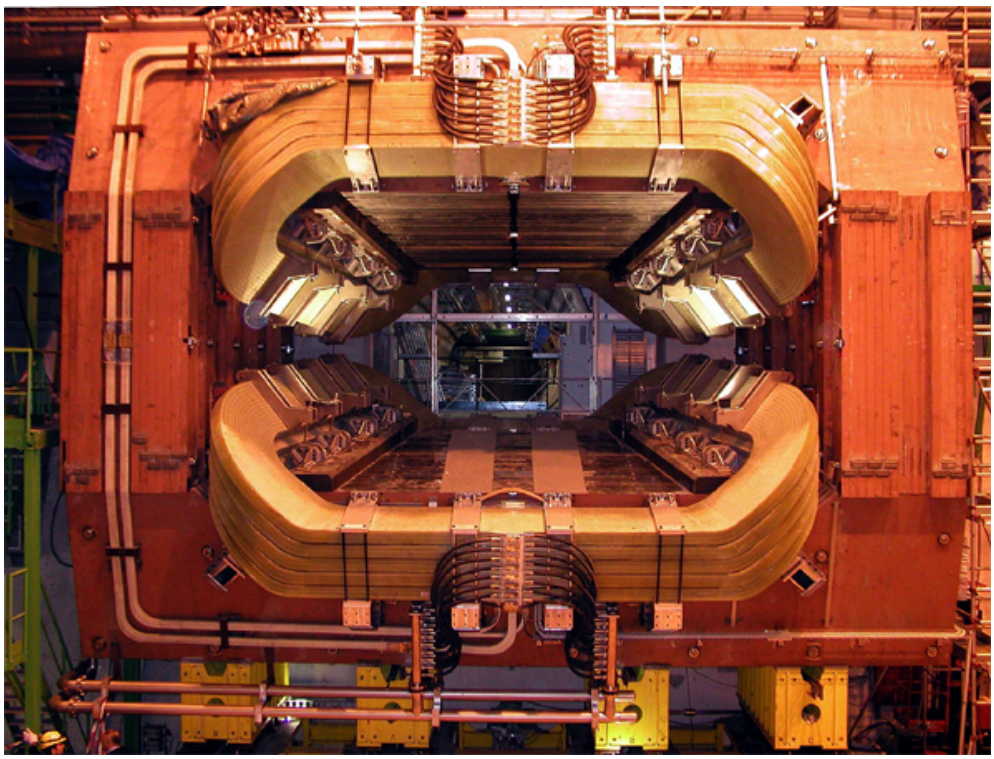

Figure 2.9: *The LHCb magnet. The magnet weighs 1600 tons and measures 11 m* $\times 8$ *m* $\times$ 5 *m.*

In order to achieve the required momentum resolution for charged particles, the magnetic field must be measured with a relative precision of a few times 10^{-4}. To measure the magnetic field integral $\int Bdl$ with this precision, as well as the position of the B-field peak with a precision of a few millimeters, an array of Hall probes is used. The Hall probe array consists of 60 sensor cards mounted on a G10 support covering a grid of 80 mm $\times$ 80 mm. Each sensor contains three probes that are mounted orthogonally on a cube, together with a temperature sensor and the read-out electronics.

The magnetic field has been measured using a grid of $8 \times 8 \times 10$ cm^3 with a precision of about $4{\cdot}10^{-4}$ in the full tracking area of the LHCb spectrometer [35], extending from the interaction point to the RICH2 detector and covering most of the LHCb acceptance region. Assuming the procedure of demagnetization of the iron yoke is carried out correctly, the absolute field value is found to be reproducible for both polarities with a precision better than $4{\cdot}10^{-4}$.

The magnetic dipole field is reversed frequently to allow the cross-check of systematic left-right effects in the detector.

2.3.4 Inner Tracker

The Inner Tracker (IT) silicon strip detector is located behind the LHCb magnet. The silicon sensors are distributed in four stereo layers in an orientation similar to the TT: the first and fourth layer are oriented vertically, the second is oriented with an angle of $-5°$ around the z-axis, and the third is oriented with an angle of $+5°$. Each sensor has 384 strips and covers an area of 7.6 cm $\times$ 11.0 cm, with a sensor thickness of 320 μm for ladders with single sensors and 410 μm for ladders consisting of two sensors. The single-sensor ladders are located below and above the beam pipe, while the double-sensor ladders are on the sides. The thickness of the sensors is chosen such that the signal-to-noise ratio in the two cases is similar. The layout of an inner tracker layer is depicted in Fig. 2.10.

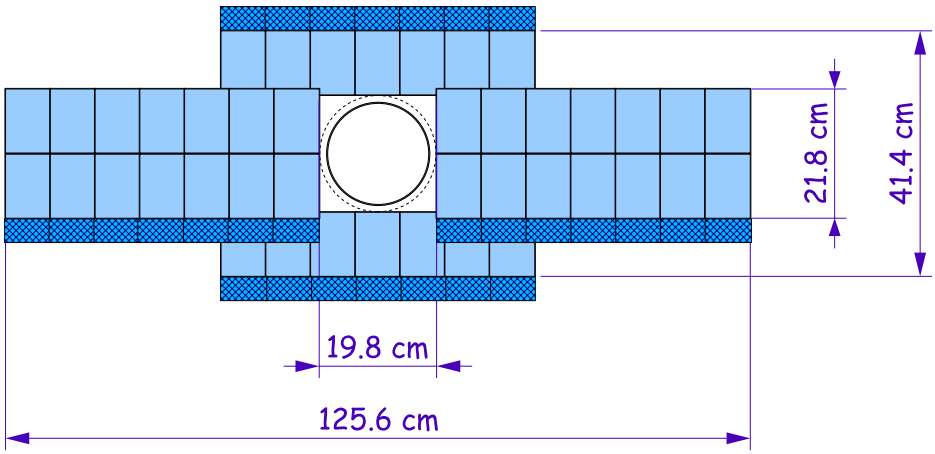

Figure 2.10: *Layout of one layer of IT sensors. The beam pipe is located in the center of the four boxes containing the sensors.*

It is important to have high granularity since, due to the forward production of B-mesons, the highest particle densities will occur in this region. With an average occupancy of about 2%, the Inner Tracker ensures good pattern recognition capabilities.

2.3.5 Outer Tracker

The Outer Tracker (OT) is located behind the dipole magnet and covers the large detection surface surrounding the Inner Tracker. It consists of three stations with four layers each, where each layer is again composed of a left and a right half containing nine detector modules of 128 straws. The Outer Tracker is a gas detector that uses a mixture of 70% Argon, 28.5% CO_2, and 1.5% O_2 as drift gas in the straws, while its walls are made of conductive material. The straws have an inner and outer diameter of 4.9 and 5 mm respectively, and are mounted adjacently to one another, with a pitch of 5.25 mm. Each straw contains an anode wire operated at a high voltage (HV) of about 1550 V. The module layout in an OT station is depicted in Fig. 2.11.

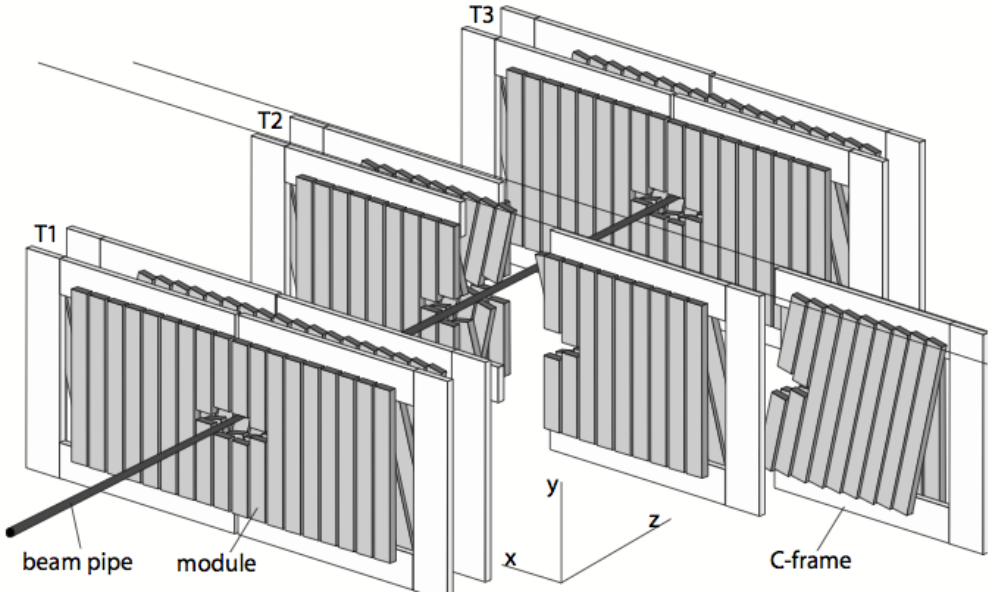

Figure 2.11: *Schematic layout of the LHCb OT detector. Each of the tracking stations T_1, T_2, T_3 accommodates four layers of detector modules. The individual 5 m long modules are mounted in so-called C-frames.*

When a charged particle traverses a cell, the gas is ionized and the primary and secondary electrons produced in this process drift to the anodes under the influence of the electric field. Their drift times are registered by a time-to-digital converter (TDC) read-out, giving a position with a resolution of about 200 μm for the measurement of the track position [36]. The efficiency of a single straw tube depends on the ionization properties of the gas, the HV settings and the discriminator threshold in the read-out electronics. Under nominal operations the efficiency is about 99%. The average occupancy in the OT is about 13% at an instantaneous luminosity of $4 \cdot 10^{32}$ cm^{-2}s^{-1} [37].

2.3.6 Track and vertex reconstruction

Track reconstruction in LHCb is performed combining the hits of all tracking detectors. Once the track is reconstructed, the curvature of the track is combined with the magnet field map to determine the charges and momenta of particles. The track reconstruction starts with pattern recognition algorithms that find the *track seeds* in the VELO or T-stations. Subsequently these seeds are combined to form tracks. The tracks are then fitted using a Kalman filter method [38, 39], in a way such that measurements are progressively added in the fit to update the track

state. In this fit method, effects of multiple scattering and energy loss are taken into account. There are five types of tracks defined at LHCb:

- **Long tracks** - These tracks are reconstructed using measurements from all tracking stations; they therefore have the best momentum resolution.
- **Upstream tracks** - These are low momentum tracks that traverse the VELO and TT before being bent out of the detector acceptance by the magnetic field.
- **Downstream tracks** - These tracks are typically the decay products of particles that either decay outside the VELO acceptance or do not have enough hits in the VELO. To reconstruct these tracks, measurements from the TT and T-stations are used.
- **T tracks** - These are either tracks produced in secondary interactions or small angle tracks, and are measured only in the T-stations.
- **VELO tracks** - These are large angle or backward produced tracks that are used mainly for primary vertex reconstruction. Due to their large or backward production angle they only have hits in the VELO detector.

A typical B event includes on average about 100 reconstructed tracks, the majority of which are long tracks or VELO tracks.

The information from long, upstream and VELO-tracks is also used to reconstruct primary vertices, by using the information on the impact parameter (IP) of a track with respect to the interaction vertex. The position of the vertex is determined by minimizing a global least square fit:

$$\chi^2 = \sum_{i=1}^{n_{tracks}} \frac{d_{0i}^2}{\sigma_{d_{0i}}^2}$$

where d_{0_i} is the three-dimensional IP between a track i and the primary vertex position and σ_{0_i} is the track impact parameter resolution. Secondary (decay) vertices instead are reconstructed offline using fully reconstructed B decays. Further information on these techniques can be found in [40].

2.4 Particle Identification System

Particle identification in LHCb is provided by three different types of detectors: the Ring Imaging Cherenkov detectors (RICH1 and RICH2), the calorimeter system and the muon system.

2.4.1 Cherenkov detectors

According to the Cherenkov effect, a charged particle transversing a dielectric medium at a velocity higher than the speed of light in that medium will emit radiation in a cone at a specific angle:

$$cos(\theta_c) = \frac{1}{\beta n} \tag{2.1}$$

where θ_c is the Cherenkov angle, β the ratio between the particle's velocity and the speed of light and n the refractive index of the material. The charged particle will only emit radiation if $\beta \geq 1/n$. If the velocity of the particles is too high then the Cherenkov angle saturates; all particles will radiate at the same angle and particle identification is no longer possible.

The two detectors use different gases in order to cover a different momentum range for identification: RICH1 uses silica aerogel and C_4F_{10}, and RICH2 uses CF_4. The RICH1 detector allows for 3σ K/π separation from 2 GeV up to 50 GeV, while RICH2 provides 3σ

K/π separation up to 100 GeV.[2] The RICH1 covers the full angular acceptance of 300 mrad (horizontal)×250 mrad (vertical), while the angular acceptance of the RICH2 only corresponds to 120 mrad (horizontal)×100 mrad (vertical), covering the area where most high momentum particles are produced. The layout of the RICH detectors is shown in Fig. 2.12.

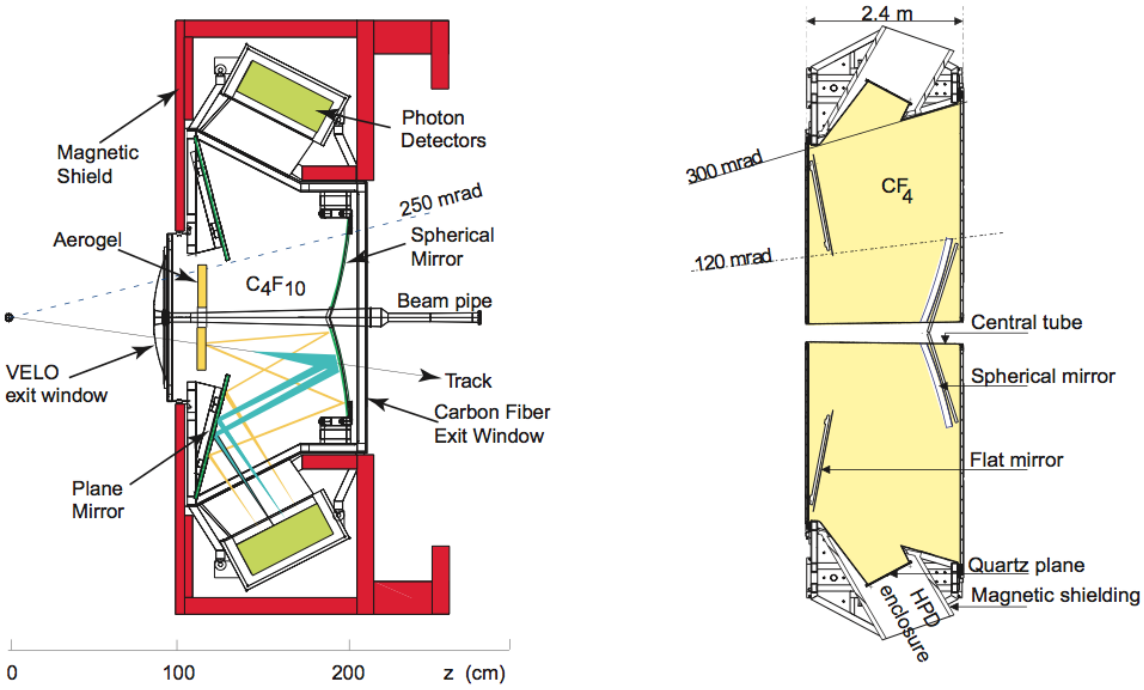

Figure 2.12: *Left: side view of the RICH1, located in front of the magnet. Right: top view of RICH2, located behind the magnet.*

2.4.2 Calorimeters

The calorimeter system is used to identify electrons, photons and hadrons by measuring their energy deposition inside the calorimeter. Particles interact with the calorimeter material (photons and electrons via electromagnetic processes, hadrons via nuclear interactions) and produce a cascade of secondary particles which are absorbed, while ionizing the material. Inside the calorimeter layers of scintillator material are located in-between the layers of absorber material, and the scintillation light produced in these layers is detected using photomultiplier tubes via wavelength shifting fibers. The number of detected photons is proportional to the energy deposition in the material.

The calorimeter system consists of an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL), preceded by a pre-shower detector (PS) and a scintillator pad detector (SPD). In the SPD charged particles will produce scintillation light, which distinguishes them from neutral particles, while in the PS electrons and photons will produce a shower, allowing for the separation of electrons, photons and hadrons. The ECAL consists of 66 alternating layers of 2 mm thick lead sheets and 4 mm thick scintillator tiles, perpendicular to the beam axis. The HCAL has a similar structure, but in this case the layers are parallel to the beam axis. The calorimeter segmentation is shown in Fig. 2.13.

[2]The Kaon identification efficiency is approximately 95%.

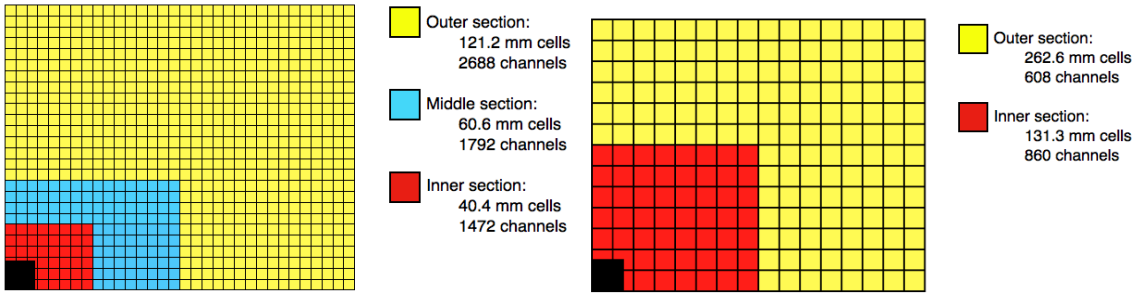

Figure 2.13: *Lateral segmentation of the ECAL (left) and HCAL (right). The lengths of the cell sides are given in the legends.*

The energy resolution that can be obtained by the two calorimeters is[3]:

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{(E/GeV)}} \oplus 1.5\% \text{ (ECAL)}$$
$$\frac{\sigma(E)}{E} = \frac{80\%}{\sqrt{(E/GeV)}} \oplus 10\% \text{ (HCAL)}$$

2.4.3 Muon System

Muon identification is important in LHCb because muons are present in the final states of several CP-violating B decays and because muons of semi-leptonic B decays are used as a flavour tag.

The muon system consists of five separate tracking stations located at the downstream end of the spectrometer; one in front, and four behind the calorimeter system. The layout of the muon stations and the geometry of the muon tiles are shown in Fig. 2.14.

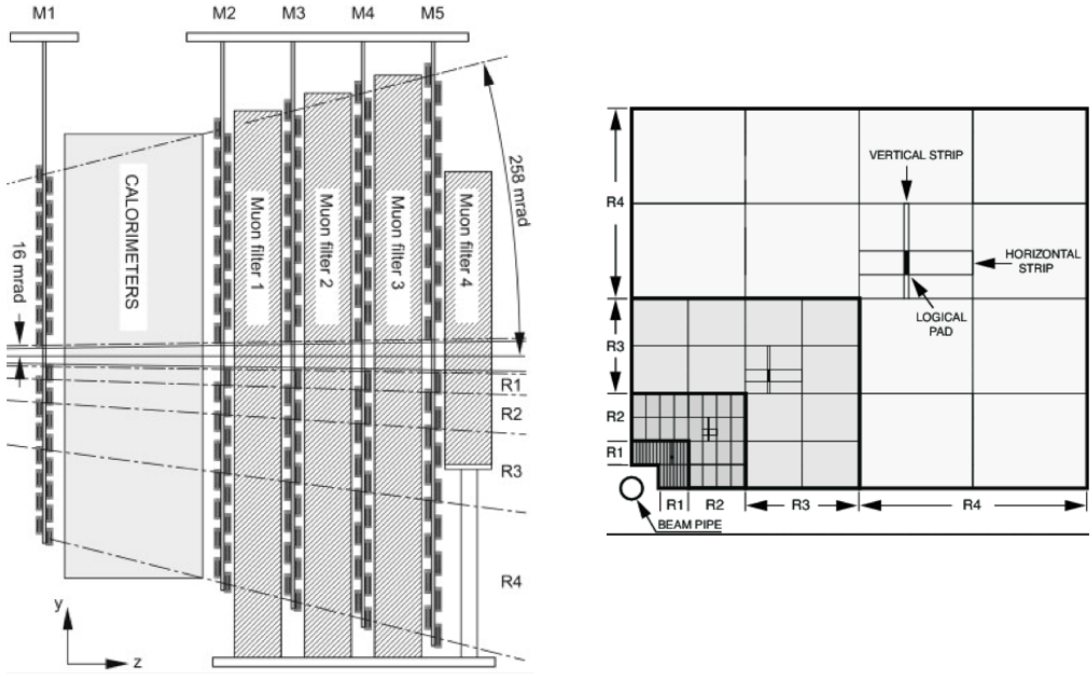

Figure 2.14: *Layout of the muon stations (left) and geometry of the muon tiles (right).*

[3]In the following expressions, the symbol $\oplus$ indicates addition in quadrature.

The four muon stations positioned behind the calorimeters are separated by 80 cm thick absorbers made of iron, which provide the required attenuation of hadrons.

Since the rate of the particle flux is not constant across the (x, y) plane of the muon detector, a muon station is divided in four regions with different granularity. Since the inner region of the first station M1 has the highest charged particle densities and because of aging considerations, triple gas-electron multiplier (GEM) detectors are used in this region. In the remaining regions, muon detection is provided by multi-wire proportional chambers (MWPCs). The segmentation in regions, sectors, channels and logical pads is optimized according to the particle fluxes. The spatial resolution is better in the x direction than in the y direction, which is used for momentum measurements in the bending plane. A fast counting gas ($Ar/CO_2/CF_4$) is employed since the muon system is used in the L0 trigger, meaning that a fast response of the detector is required.

2.4.4 Methods for particle identification

The information from the PID detectors is used to identify charged particles such as $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$ and p, as well as neutral particles such as γ and π^0.

The identification of charged pions and kaons is performed using the information from the RICH detectors, via an algorithm that matches the observed pattern of Cherenkov radiation in the photodetectors with the pattern expected under pion or kaon hypotheses by using a global log-likelihood method. To identify muons, tracks reconstructed by the tracking system and having a momentum $p \geq 3$ GeV/c are extrapolated to the muon system. Hits associated to these tracks are then looked for in all five muon stations in a Field of Interest (FoI) built around the extrapolation position [41]. Tracks with p <3 GeV/c will not reach the muon system. Electrons are identified in the ECAL, by looking for a balance between the measured track momentum and the energy deposition in the calorimeter. In addition, the position of the cluster in the ECAL must match to the extrapolated impact point of the track. Photons are identified by looking for clusters without an associated track.

At LHCb, kaons are identified with an average efficiency of approximately 95%, with an associated average rate of pion misidentification of about 10% [42]. The efficiency for muon identification was determined by using $B_0 \to J/\psi K_s$ decays, and is equal to about 94%. Further information on the particle identification methods used at LHCb can be found in [43, 44].

2.5 Trigger System

The LHCb trigger consists of a two-stage system: the hardware L0 Trigger and the software High Level Trigger (HLT). The HLT in turn is subdivided in two stages: HLT1 and HLT2.

2.5.1 L0 Trigger

The L0 Trigger is a hardware trigger whose main goal is to reduce the event rate from the bunch crossing rate of 40 MHz down to 1.1 MHz, the maximum rate at which the detector can be read out. The L0 trigger is divided in four components: the L0 Calorimeter Trigger, the L0 Muon Trigger, the Pile-Up trigger and the L0 decision unit. Each of these configurations is built by selecting different combinations of cuts on variables such as the energy depositions in the calorimeters and muon stations, the Scintillating Pad Detector (SPD) multiplicity and the Pile-Up (PU) multiplicity.

To pass the L0 calorimeter trigger, either the transverse energy (E_T) of the HCAL cluster must be larger than 3.5 GeV or the E_T of the ECAL cluster must be larger than 2.5 GeV. In the muon system, the transverse momentum (p_T) of particles is reconstructed by forming a track in all five stations and assuming that the track originated in the nominal interaction point $(x, y, z) = (0, 0, 0)$. The L0 muon trigger requires the transverse momentum (p_T) of a single

muon to be at least 1.48 GeV/c. An additional di-muon requirement is implemented, for which the product $p_{T_1} \cdot p_{T_2}$ must be larger than 1.68 $\mathrm{GeV}^2/\mathrm{c}^2$. If one or more of the above requirements are satisfied then the L0 decision is positive, and the information of all subdetectors is read out by the data acquisition system (DAQ).

In 2010, the PU multiplicity information was used in combination with other variables in the L0 Minimum Bias configuration. Since 2011, the PU information is used to select Beam-Gas events using the fact that the PU observes tracks in the backward direction [45]. Moreover, in 2012 the PU multiplicity information was used in the selection strategy to study Central Exclusive Production (CEP) at LHCb (see [46]). In particular, the selection of CEP events with hadronic decay products was improved by introducing a veto on activity in the backward direction.

2.5.2 High Level Trigger

Events that are accepted by the L0 trigger pass to the second stage, the High Level Trigger (HLT), at a maximum rate of 1.1 MHz. The HLT consists of two stages, HLT1 and HLT2, which run on a computing farm with approximately 16,000 cores.

The HLT1 performs a partial event reconstruction in the region around the object that caused the L0 to trigger, in order to confirm or reject the decision of the L0 trigger. Particles in the VELO are reconstructed and the positions of the primary vertices (PV) in the event are determined. In order to limit the CPU usage VELO tracks are selected based on the quality of the tracks and requiring a minimal IP with the PV. For these VELO tracks, long tracks are formed (forward tracking) such that momentum information is available. The HLT1 reduces the rate from 1.1 MHz to 50 kHz, which is low enough to then allow forward tracking of all VELO tracks.

Subsequently, HLT2 searches for secondary vertices and performs complete pattern recognition and a full event reconstruction, selecting either distinct decay modes or specific event topologies. The output rate is reduced from 50 kHz to 4 kHz such that events can be written to storage and processed offline.

In addition, LHCb employs a novel *deferred triggering* technique in its HLT farm. Thanks to this deferred HLT, about 20% of the events can be stored at the input of the HLT and processed during the inter-fill periods of the LHC machine. The deferred triggering increases the efficiency of the LHCb trigger system, and causes the output rate of the HLT to increase from 4 kHz to 5 kHz.

Chapter 3

Error Monitoring in the VELO Detector

In order to reconstruct particle tracks without any bias in the VELO detector, it is important that the digital representation of the data accurately represents the signal charge deposited in the sensors and collected at the silicon strips. Sources of possible errors may occur at different stages of the data acquisition chain, for example when the signals are transferred to the preamplifier or when the data is transferred from the detector to the Data AcQuisition system (DAQ). A study of individual sources of error is performed in order to eliminate them or keep any remaining ones at an acceptable level.

In this chapter the VELO detector is described in more detail with emphasis on the read-out system. An error monitoring tool is included in the VELO read-out system, which reports in so-called error banks. Different sources of possible errors in the readout and transmission of data in the VELO data acquisition chain are studied. In Section 3.1, the geometry and DAQ of the VELO detector are presented. The format of the data is described in Section 3.2, and the format of the error banks is given in Section 3.3. In Section 3.4 the different types of errors that cause the generation of an error bank are discussed in more detail, and in Section 3.5 the content of the error monitoring is illustrated.

3.1 The VELO Detector

3.1.1 Detector modules and silicon sensors

The VELO is a silicon strip detector made up of two moveable halves, located in vacuum around the interaction region of the LHC beams [34]. The VELO halves are designed to 'open' and 'close' before LHCb beam injection and data taking, respectively. When open, the sensors are retracted by 29 mm from the center, in order to avoid beam-induced damage to the sensors. During physics data collection the two halves are closed to form a circular acceptance, and the sensitive area of the detector then starts at 8.2 mm from the beam. The closing of the two VELO halves is a fully automated procedure, performed by an application called the Closing Manager, implemented in the control software package PVSS, and a programmable logic controller (PLC), which controls the stepping motors. The two detector halves close in five steps: 58 mm, 40 mm, 20 mm, 10 mm, 2 mm separation between the two halves, to the 'fully closed' position, when the active region of the sensors is at 8.2 mm from the beam. At each step the distance from the luminous region is calculated by reconstructing the vertices of the pp-interactions. Two sets of 400 vertices, one for each detector half, are reconstructed and for both the distance to the beam is calculated. If the two measurements are in agreement and various other safety conditions are met, then the detector is moved closer to the beam. The entire closing procedure takes 3.5

minutes.

Each VELO half has 42 VELO modules with two silicon sensors mounted back-to-back, one with R-strips and one with ϕ-strips. In the upstream region 2 Pile-Up stations consisting only of R-sensors are placed. How the sensors are arranged in the vertex detector is depicted in Fig. 3.1.

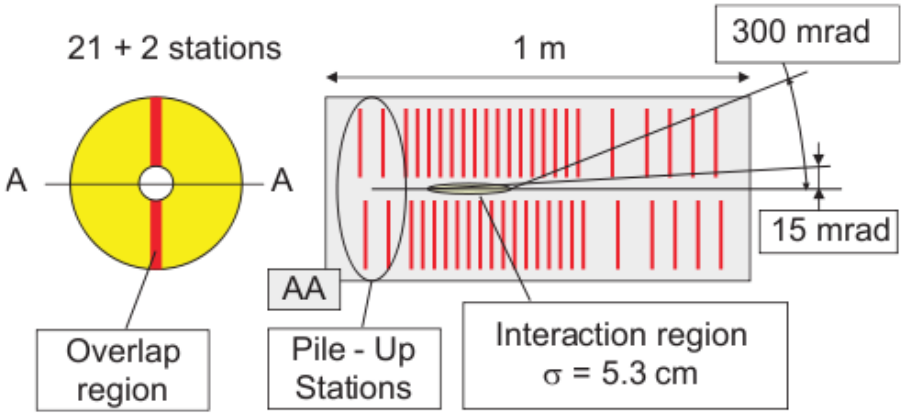

Figure 3.1: *A schematic drawing showing how sensors are positioned in the VELO detector.*

The geometry of the silicon strips is such that the pitch is smallest at low radius (see Sec. 2.3.1) and is therefore optimized to give the highest vertex resolution for tracks close to the inner region. In Fig. 3.2 an enlarged view of an R and a ϕ sensor are shown.

The ϕ sensors are divided in an inner and an outer region (not well visible in the picture), which has twice as many strips as the inner region. The signal routing lines of the inner strips pass over every other strip of the outer region before reaching the readout electronics. As a consequence, half of the strips of the outer region have an overlaid routing line, an aspect that is important in a later analysis described in Ch. 5.

When a charged particle traverses the detector it ionizes the material, generating electron-hole pairs. Due to the High Voltage (HV) the electrons then drift to the strips and form the signals, which are amplified, shaped and stored. The sensors are 300 μm thick and use n-on-n technology, with an active region that goes from 8.2 mm to 42 mm from the center. Only one station upstream with respect to the interaction region uses n-on-p technology. The sensors of this module were mounted in order to study their performance in a high radiation environment.

The silicon sensors are kept at a temperature of $-8°$C during operation, in order to reduce the effects of radiation damage and the risk of thermal runaway due to an increase in the bias current. The temperature of the silicon sensors is maintained by a CO_2 cooling system, which is operated at an evaporation temperature of $-28°$C. The cooling system has been working reliably since the start of operation in 2009.

The sensors are placed in a secondary vacuum with respect to the beam volume, shielded by a 300 μm thick aluminum box, known as the RF-foil. The RF-foil has an additional purpose: it shields the modules from the interference of the beams as it conducts the mirror currents in the interaction region of LHCb.

3.1.2 Data acquisition system

The differential signals that are to be received or sent by the electronic circuits on the modules are transmitted by repeater boards that are located directly outside of the vacuum tank. There

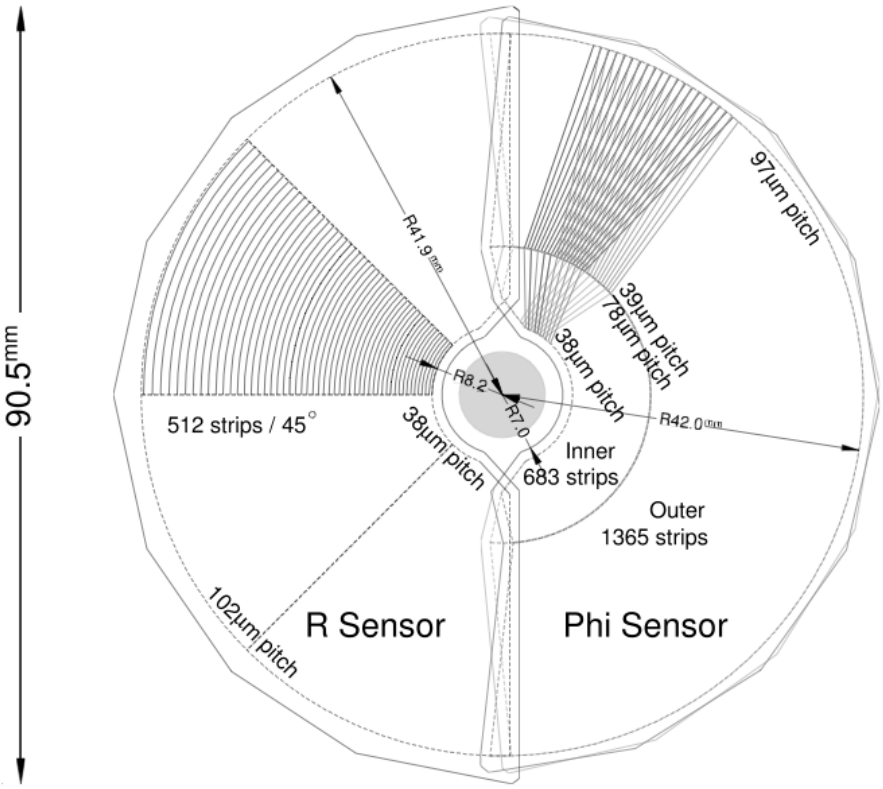

Figure 3.2: *Schematic layout of an R- and $\phi-$sensor in which the strip geometries are depicted. On the left, an R-sensor is shown. On the right, two adjacent ϕ-sensors are shown superimposed to show the effect of the stereo angle of the ϕ-strips.*

is one repeater board for each sensor. The analogue signals coming from the repeater boards are sent over 60 m copper cables to the TELL1 (Trigger ELectronics Level 1) boards [47] of the data acquisition system, which are readout boards used for complex processing of raw data and zero suppression. To read out the 21 VELO and 2 Pile-Up stations a total of 88 TELL1 boards are used. A TELL1 consists of a motherboard hosting four analogue receiver (ARx) cards and five Field Programmable Gate Arrays (FPGA); four Pre-Processing FPGAs (PP-FPGA) and a "Synchronization and Link" FPGA (SyncLink-FPGA), and a Front-end EMulator (FEM) mezzanine [48]. The readout path between a VELO sensor and a TELL1 is shown in Fig. 3.3.

Each sensor is connected to 16 Beetle chips, 64 links with 32 channels per link, and one TELL1 board. The ARx cards receive the analogue signals, where the digitization of the analogue signal is done with 10 bit precision, converting the analogue signal of each Beetle channel to a value in the range $0-1023$ ADC counts [49]. There are four ARx cards on each TELL1 board, each of which has an input capacity of 16 analogue links. The data that arrive at the TELL1 is processed by the four PP-FPGAs, for example, performing the pedestal calculation, common mode suppression and zero-suppression of data. The SyncLink-FPGA is in charge of collecting and merging the data fragments from the PP-FPGAs and distributes synchronization signals to these pre-processors. Further details on the data processing steps and on the processing algorithms implemented in the five FPGAs can be found in [50]. The Front-end Emulator (FEM) is a Beetle front-end chip mounted on a mezzanine card. It is connected with the clock, reset and L0 trigger signals and is used for synchronization purposes such as the Pipeline Column Number (PCN) check (see Sec. 3.2), where the PCN gives the position in the buffer at which

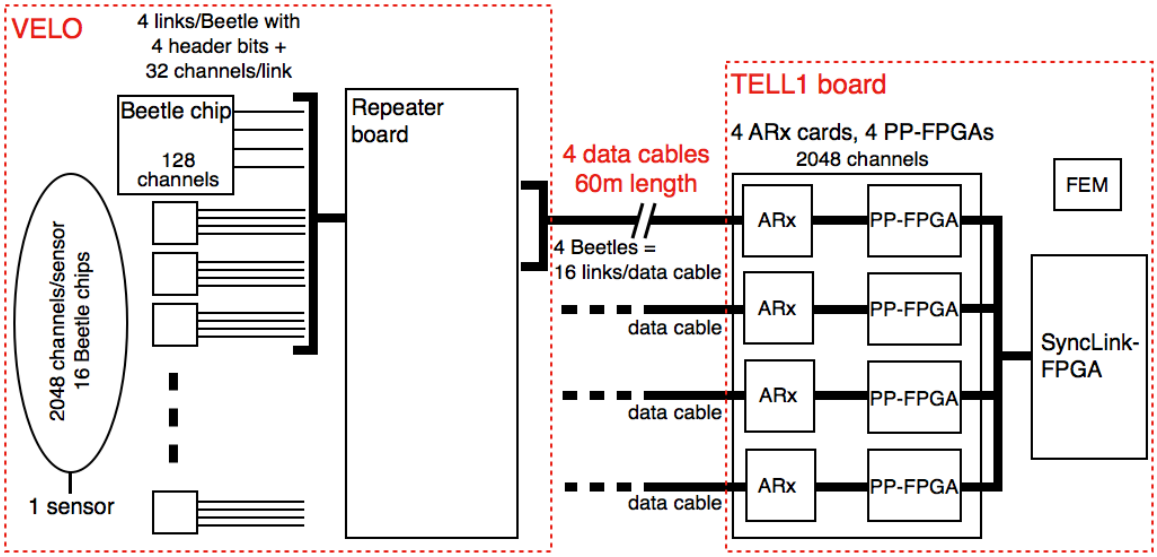

Figure 3.3: *Read-out path between a VELO sensor and a TELL1 board.*

the data is stored. The FEM PCN and the PCN of the other Beetle chips connected to that TELL1 are compared in the SyncLink-FPGA. If their values are not equal, the corresponding PCN error bits are set and an error bank is generated.

The FEM is also responsible for setting the *ADC FIFO error* word in the error bank in case something goes wrong with the actual data transmission. A FIFO (First In First Out) is a buffer that consists of a set of read and write pointers, storage and control logic. At initialization both the read and write pointers point to the beginning of the buffer. In the VELO the FIFO is synchronous, which means that the same clock is used for reading and writing. With every read/write operation the corresponding pointer is incremented, with the condition that the read pointer cannot overtake the write pointer. If the read pointer reaches the write pointer the FIFO signals empty and the EMPTY flag is set. If the write pointer reaches the read pointer the FIFO sets the FULL flag. In case this mechanism does not work properly an error, known as *ADC FIFO error*, is generated. The other types of possible errors that can occur at different stages of the data acquisition chain are discussed in the following sections.

The readout system of the VELO is shown in Fig. 3.3. The data of the total of 88 silicon sensors[1] is collected in events that are structured in blocks corresponding to TELL1 units. To each TELL1 unit corresponds one silicon sensor. The data of each sensor is read out by 16 Beetle chips having 4 A-Links each. Each A-Link contains the data of 32 of the 2048 silicon strips on a sensor, preceded by 4 additional header bits. A schematic overview of how the data of an event is structured is shown in Fig. 3.4 for one TELL1 unit.

3.1.3 The Beetle front-end readout chip

Each silicon sensor is connected to 16 Beetle read-out chips, each having 128 channels and designed to operate at 40 MHz. The input signals are transferred to charge (pre)amplifiers that act as charge-to-voltage converters, producing an output voltage proportional to the input charge. The output voltage pulse is then shaped by consecutive steps of integrators and differentiators, and the amplitude of the output pulses is stored in a 187 cells deep pipeline buffer. A simplified functional diagram of the front-end circuit of the Beetle is shown in Fig. 3.5.

[1] 84 VELO sensors and 4 sensors for the Pile-Up.

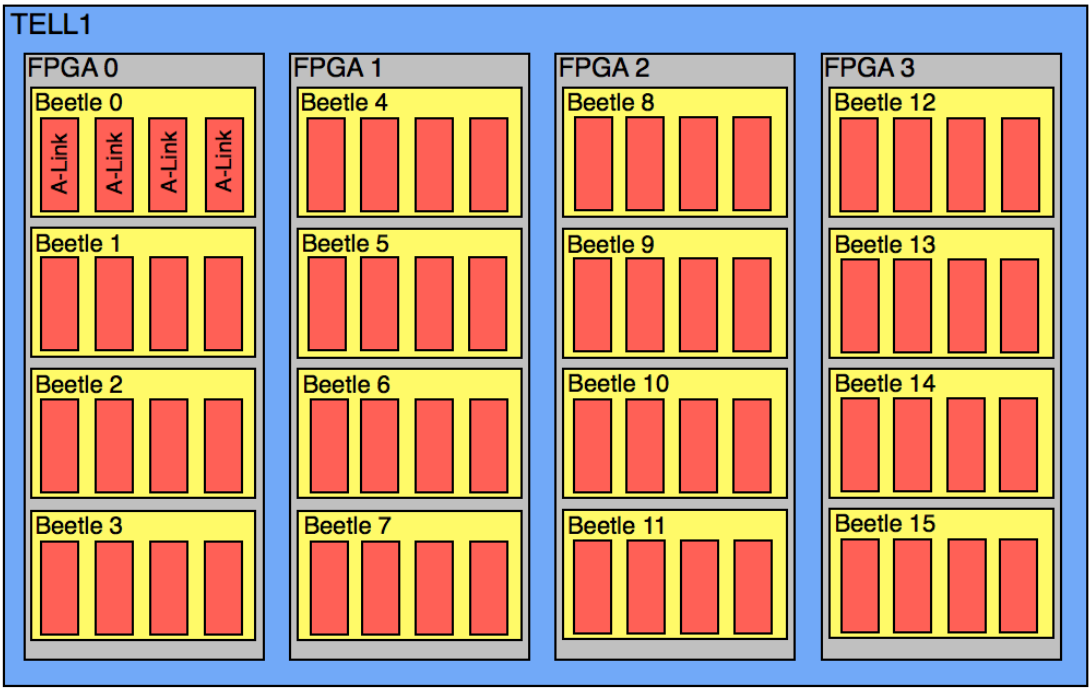

Figure 3.4: *Schematic of the structure of the data collected in one event, for one TELL1 unit. The data is divided in TELL1s, each containing 4 PP_FPGAs with again 4 Beetle chips each.*

In the pipeline buffer events can be stored for 4 μs while waiting for a trigger decision. The address of a signal in the pipeline is given by the pipeline column number (PCN). The addresses that need to be read out are stored in a 16 cells deep address buffer until the signals are retrieved from the pipeline. The trigger must be synchronized with the Beetle, in order to read out the correct pipeline position.

After receiving an L0 trigger, the output of the 128 channels of the Beetle is multiplexed onto 4 ports (the so-called A-Links). Each port carries the information of 4 header bits, that flag information on the data, and 32 analogue channels, containing the observed silicon strip signals [51], as shown in Fig. 3.6. The transmission of data takes (32 *analogue channels* + 4 *header bits*) $\times$ 25 ns = 900 ns per trigger, allowing for continuous readout at 1.1 MHz as required by the Level 1 trigger.

Of the four header bits that are present for each port, the first two are IHeader bits and the last two give the PCN. The meaning of the various header bits is summarized in Table 2.1.

In the table, the IHeader bits are labeled $I_0 - I_5$, S_0 and S_1, while the PCN bits are labeled $P_0 - P_7$. Most of the I_n header bits correspond to *parity bits*; bits that act as a check on a set of binary values. An (even) parity bit is calculated such that the number of '1's in the set plus the parity bit must always be even. The I_3 and I_4 header bits refer to the comparator circuits consisting of an integrator, a threshold generator and a discriminator, which convert the input signals from analogue to binary values.

The definition of the IHeader bits for Beetle 1.5 is as follows:

- I_0 : *LeadingBit* (this is always equal to 1).
- I_1 : *ParTpSelect*, parity of 128-bit Test-pulse Selection register (TpSelect). At the moment this register is not active (i.e., bit is set to 0).

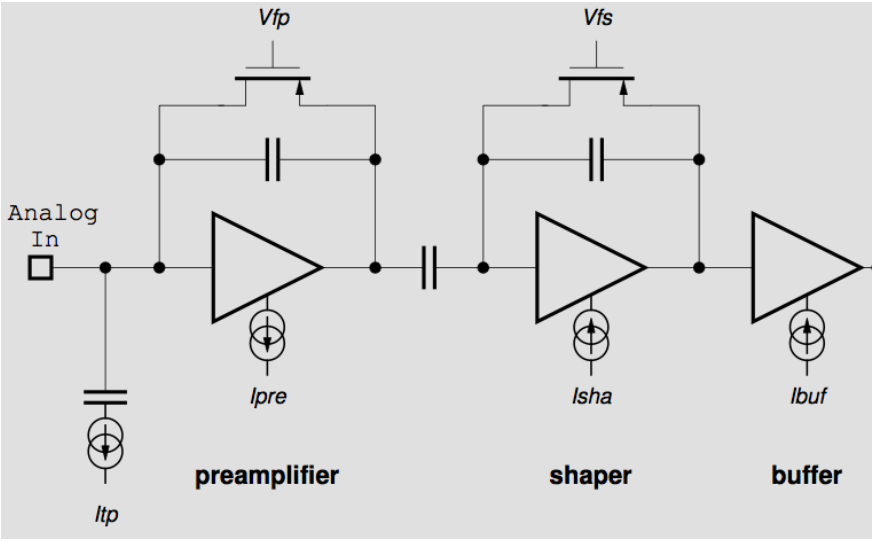

Figure 3.5: *Front-end circuit of the Beetle chip. The charge pre-amplifier, shaper and pipeline buffer are shown.*

Bit		Description
I_0	LeadingBit	always active (=1)
I_1	ParTpSelect	(even) parity of register *TpSelect*
I_2	ActiveEDC	1 indicates active error detection and correction (EDC) logic
I_3	ParCompChTh	(even) parity of register *CompChTh*
I_4	ParCompMask	(even) parity of register *CompMask*
I_5	ParPCN	(even) parity of pipeline column number (PCN)
S_0		LSB of register *SEUcounter*
S_1		bit 1 of register *SEUcounter*
P_0		LSB of pipeline column number
P_1		bit 1 of pipeline column number
P_2		bit 2 of pipeline column number
P_3		bit 3 of pipeline column number
P_4		bit 4 of pipeline column number
P_5		bit 5 of pipeline column number
P_6		bit 6 of pipeline column number
P_7		MSB of pipeline column number

Table 3.1: *Definition of header bits for the Beetle 1.5 readout chip used in the VELO.*

- I_2 : *ActiveEDC*, if I_2 is equal to 1, it indicates the detection and correction of an error. This error corresponds to a Single Event Upset (SEU), which is the change of state of a memory register induced by ionization of a traversing particle in the chip.

- I_3 : *ParCompChTh*, parity of the comparator threshold register (CompChTh). This bit is only active for the Pile-Up detector, while for the VELO it is set to 0. The comparator threshold register (*CompChTh*) selects the number of delta thresholds to subtract from the global threshold. It is operated as a shift register: first, the delta thresholds are calculated for each channel by programming the *CompChTh* 128 times consecutively, then, the bits are assigned to the different channels by a shift mechanism.

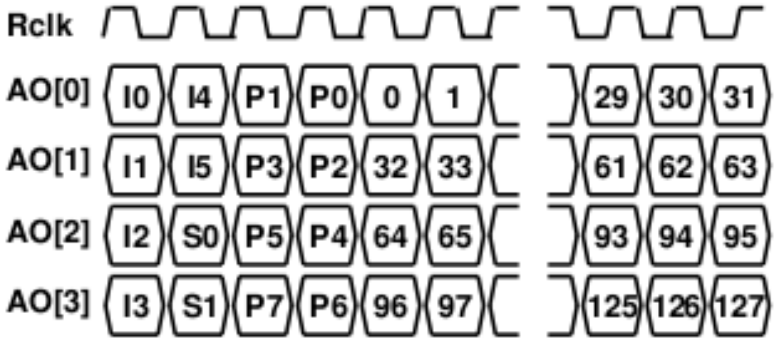

Figure 3.6: *Analogue readout mode: 128 analogue channels are multiplexed onto 4 ports, AO[0]-AO[3]. In each port the first four clocks cycles give the header bits, and they are followed by 32 analogue strip readout values.*

- I_4 : *ParCompMask*, parity of the comparator mask register (CompMask). This bit is also only active for the Pile-Up detector. The comparator mask register (*CompMask*) deactivates the operation of a single comparator channel.
- I_5 : *ParPCN*, (even) parity of the PCN.
- S_0, S_1 : Least Significant Bit (LSB) and Bit1 of the 8-bit SEU Counter, which counts the number of single event upsets that occurred in the configuration registers of the chip.

3.2 Raw data format

Three types of data format are produced by the TELL1 module: non-zero suppressed (NZS) data, zero-suppressed (ZS) data and error banks. The NZS data are used mostly for calibration or tuning purposes, while the ZS banks are used in standard data taking. The error banks flag any mistakes that have been detected by the system. These data banks are sent by the TELL1 to the DAQ. The NZS data banks contain the digitized information from all strips on a sensor, while the ZS data banks only contain information from strips that have signal above threshold, as well as combined information in the form of clusters. NZS events are typically 85 times larger than ZS events and can therefore only be collected at low rates due to bandwidth limitations. The format of the error banks will be defined and treated in more detail in Section 3.3.

In Fig. 3.7 a schematic representation of non-zero suppressed data for one Beetle chip is shown [52]. For each Beetle chip the data is arranged in 4 blocks, one for each A-Link. Each block consists of 36 10-bit words: the first 4 correspond to the headers and the remaining 32 to the analogue data of each channel.

The exact contents and format of NZS and ZS data is presented in Appendix A.

3.3 Description of error banks

The error bank is one of the data banks that can be sent by the TELL1 to the DAQ. It is generated when internal synchronization errors known as Pseudo header errors, PCN errors and

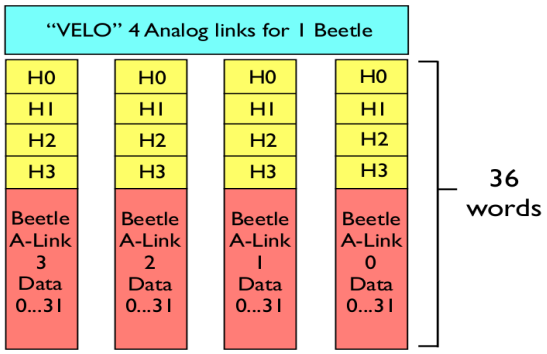

Figure 3.7: *VELO NZS data format received from the ARx cards, for one Beetle chip. For all headers and analogue channels 10 bits are used.*

ADC FIFO errors are detected, and contains the complete error information. Synchronization errors can be detected at two different stages during the data processing. The first place is the input synchronizer circuit on the PP-FPGA and the second is at the linker stage of the SyncLink-FPGA, where the event fragments from the different PP-FPGAs are merged.

3.3.1 Error bank data format

An error bank consists of fixed and optional words [53]. A ZS-event has a minimum length of 28 words and a maximum length of 52 words. In case no errors occur there are 7 fixed words per PP-FPGA, such that with 4 PP-FPGAs per TELL1 the total number of words is 28 (depicted in purple in Fig. 3.8). In case errors are present, there are 5 optional words (shown in blue in Fig. 3.8) for each PP-FPGA. Moreover, an additional word can be attached per PP-FPGA in case the data bank is either an NZS or a pedestal bank, resulting in a total of 52 words.

The details of how these error banks are decoded can be found in Appendix B.

Fixed data fields

The first word contains the event info ID, the bank list, the LHC bunch counter and the detector ID ('0x1' for the VELO). The event info ID indicates which type of trigger was used during data taking (a detailed definition of the various bits can be found in Appendix A). The 8-bit bank list defines the type of data bank. Specifically, if Bit3 of the bank list = '1', then the optional NZS bank control word is added, if Bit4 = '1' the optional pedestal bank control word is added. The second word is the L0-evtID, which is a sequential event counter. The third word is needed to cope with the variable lengths of the cluster and ADC sections. The ADC data section length gives the number of bytes of the TELL1 internal bank. In the fourth word information on which processing has been applied to the data is encoded, together with the value of the FEM PCN (see Sec. 3.1) and the number of clusters contained in the bank. The number of clusters generated for one PP-FPGA and the PCN value from the front-end emulator Beetle are also attached. The FEM PCN has the same value for the 4 PP-FPGAs. The fifth word specifies the error bank length in bytes. It is set to zero if there is no error information for a specific PP-FPGA. If the error information is attached, it is set to '0x14'. For each bank present, five optional words are attached for each PP-FPGA from which the error information was sent.

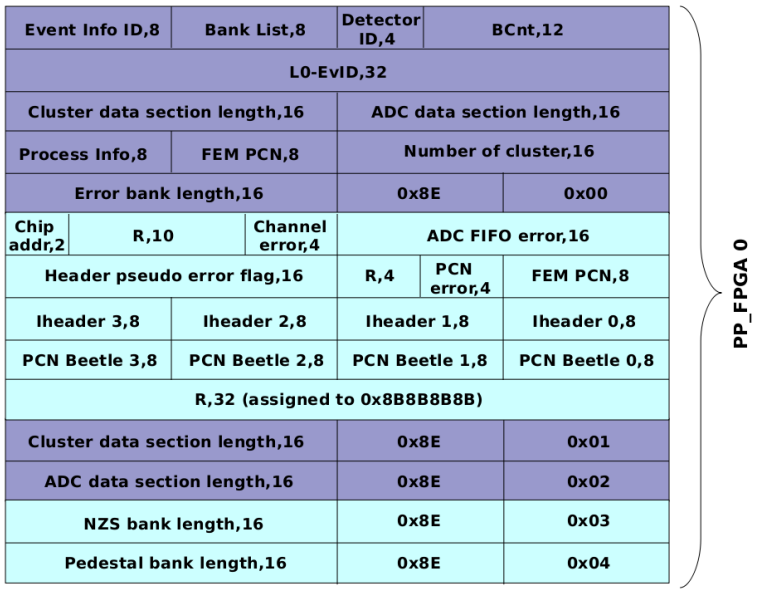

Figure 3.8: *Error bank data format. The fixed words are indicated in purple and the optional words in blue.*

Optional fields

In the optional fields the error information, NZS bank length or the Pedestal bank length are attached. The PP-FPGA info consists of the last 5 of the 8 words of the "Event Info" block. These words are attached to the data when an error is detected in the stage in which the event info is built. They contain information on possible ADC Fifo errors, PCN errors, Pseudo header errors and channel errors, as well as the values of the IHeader and PCN for the different Beetles, which will be discussed in detail in the following section.

3.4 Error Types

In this section the individual types of readout errors causing the generation of an error bank are described.

3.4.1 Pseudo header error flag

A header bit contains digital information, i.e. it is either '0' or '1'. Since the value is transmitted via an analogue signal, a high and low threshold is defined to distinguish between the two possibilities. These header thresholds are determined from the analogue header distributions. Signals that are lower than the low header threshold are digitized as '1' and referred to as *low headers*, while signals that are larger than the high header threshold are digitized as '0' and known as *high headers*. Until March 2011 the header thresholds were set per sensor, implying that the same high and low thresholds were set for groups of 64 A-Links.

In Fig. 3.9 the analogue signal of the 4 headers and 32 channels for two links are shown. These signals are centered around a pedestal value but there is a spread in the data that is due to the noise. It can be seen from the ADC values of the channels in this link that there are no

particles present in this data: it is the noise that creates a spread of the ADC values around the baseline. A typical hit (1 MIP) corresponds to approximately 40 ADC counts, so in case a silicon strip is hit we expect a signal on the corresponding channel of at least 40 ADC counts above the pedestal value. The signal of a dead link is also shown in Fig. 3.9 for comparison.

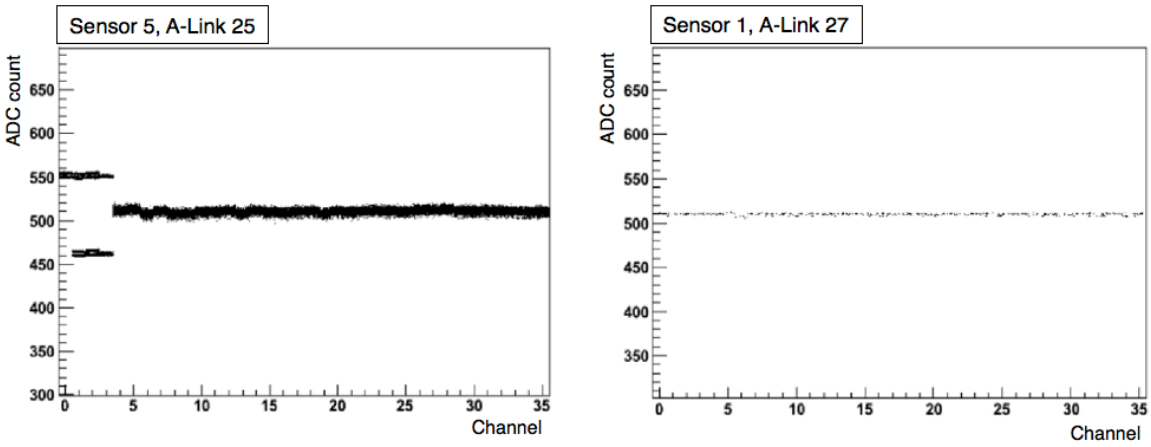

Figure 3.9: *Scatter plot of ADC value vs channel for one A-Link. Left: Header values and analogue signal for 32 channels of a correctly functioning A-Link. Right: Analogue signal for a dead A-Link is shown.*

In Fig. 3.10 the ADC distribution of all 256 header bits associated to one TELL1 are shown.

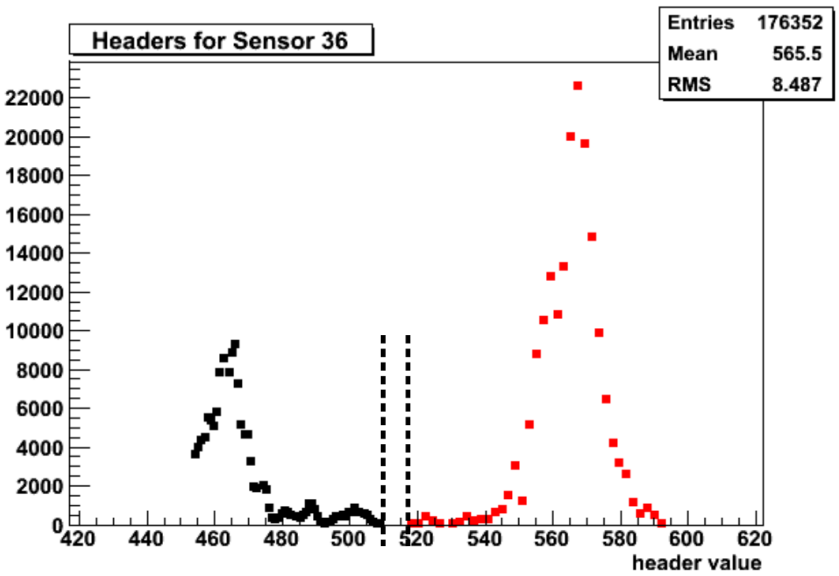

Figure 3.10: *ADC spectrum of all 256 header bits of the 64 A-Links of one TELL1. The dashed lines indicate where the high and low thresholds would be set, resulting in the high (red) and low (black) header amplitude distribution.*

If the value of a header is above the high threshold or below the low threshold the pseudo header error flag is set to '0'. If the value of a header falls in between these thresholds then the header

error flag is set to '1'. The high and low header distributions are well separated, but this is not always the case. The signals of the 64 A-Links have different offsets, which can cause an overlap of the distributions of the individual links. A better separation of the high and low header distributions is obtained when each link is studied individually.

Pseudo header errors can appear for reasons such as not optimal gain calibration and timing. The gain corresponds to the amplification the signal on a strip undergoes when it is read out by the Beetle chip and subsequently in the ARx-card. The gain changes over time and will also affect the header heights. A change in the gain can increase (decrease) the number of ADC counts corresponding to the signal, causing the values of the headers to fall in-between the header thresholds that were previously determined. The header thresholds therefore have to be recalculated with each gain calibration.

A difference between the timing of the readout and the clock on the TELL1 can also induce a pseudo header error. If there is a wrong delay in the timing the signal of the first or last channel of an A-Link can be interpreted as a header. In case no hits are present the ADC value of this signal will most likely fall in between the header pseudo thresholds and the pseudo header error flag is set.

Moreover, if one of the A-Links is dead (see Fig. 3.9) then there is no signal, the header bits are not properly set and, consequently, the pseudo header error flag is set.

The values of the headers can be misinterpreted as high or low even if no pseudo header error is present. This can happen for example if one of the header bits is stuck either in the Beetle chip or in the ARx card. In this case the value of a high header could fall below the low threshold or a low header could be interpreted as high (if the header bit is stuck to '0'). Since no header values fall in-between the header thresholds, no pseudo header error will be registered.

Fig. 3.11 shows the number of pseudo header errors present per sensor and per link for two data runs before the error analysis was done.

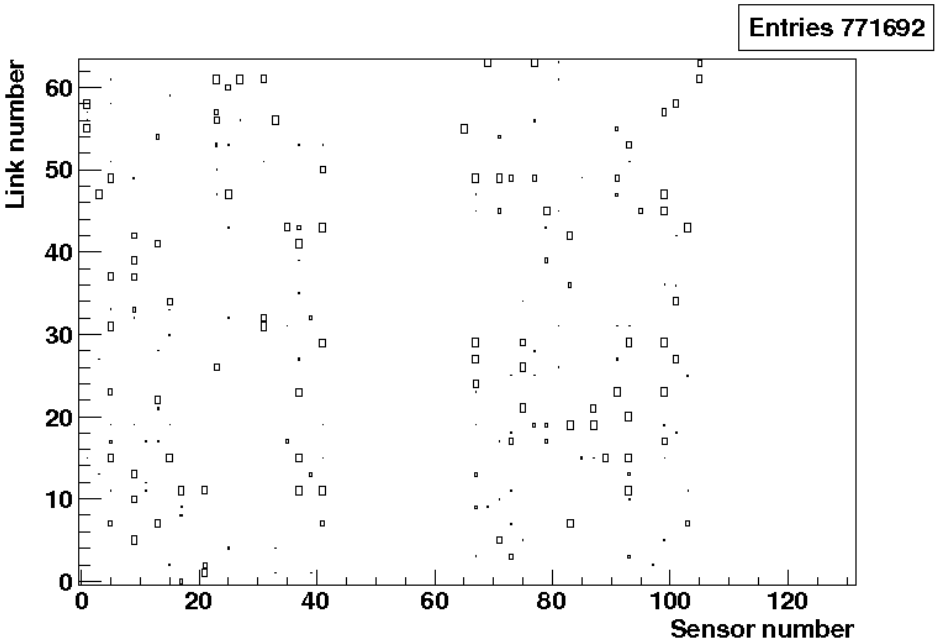

Figure 3.11: *Pseudo header errors per sensor per A-Link present at the beginning of the error analysis.*

After new values for the pseudo header thresholds were determined and uploaded, the pseudo header errors are eliminated.

3.4.2 PCN error

The Pipeline Column Number (PCN) indicates in which position of the buffer in the Beetle the data is stored. A PCN error appears if the PCN values of the sixteen Beetle chips of one sensor are not equal. As was discussed in Sec. 3.1.3, the PCN is an 8-bit number. These bits are sent out in pairs (P_0 and P_1 over A-Link 0, P_2 and P_3 over A-Link 1, etc...) over the 4 A-Links of a particular Beetle chip. The PCN value ranges from 0 to 186 and is given by:

$$P_7\ P_6\ P_5\ P_4\ P_3\ P_2\ P_1\ P_0\ ,$$

P_7 (P_0) being the MSB (LSB) [2]. If one of these A-Links is not functioning properly one or more bits of the PCN can result as 'stuck', meaning this bit is always equal to the same value ('0' or '1'). The same occurs if there are problems with the functioning of Beetle chips or ARx cards. This leads to an erroneous PCN value of that Beetle chip.

Fig. 3.12 shows the number of PCN errors per sensor per link for the situation at the beginning of the error analysis. Based on this various actions were undertaken, i.e. some defective TELL1 boards were substituted and dead links were repaired, after which most errors have been resolved.

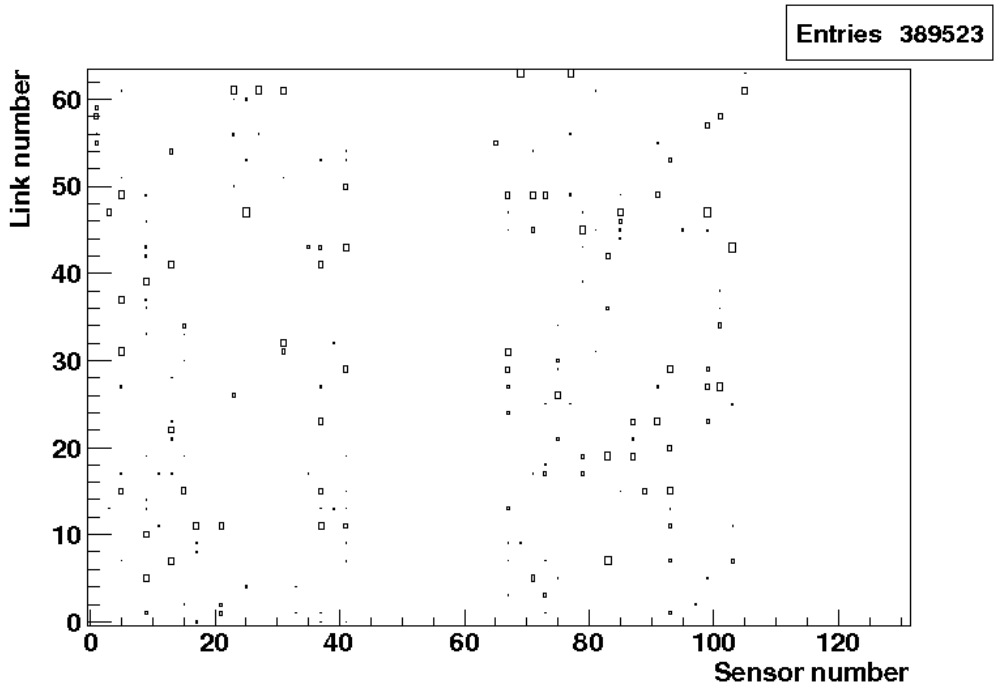

Figure 3.12: *PCN errors per sensor per A-Link at the beginning of the error analysis.*

If pseudo header errors are present for a certain A-Link, PCN errors will also appear. In fact, the presence of pseudo header errors implies that the header bits were not correctly interpreted.

[2]MSB = Most Significant Bit, LSB = Least Significant Bit

Since the last two of the four header bits contain information on the PCN, an incorrect interpretation of the header bits will likely cause the interpretation of the PCN value to be incorrect as well.

3.4.3 ADC FIFO error

The ADC FIFO is a First-In First-Out buffer responsible for the flow control of data. The Front-End-Emulator (FEM) (see Sec. 3.1) receives the trigger from the Timing and Fast Control (TFC) system and generates the data valid signal to both the SyncLink-FPGA and ARx cards. There are two conditions that can cause the ADC FIFO error flag to be set:

(i) FIFO underflow: the FEM generates a data valid signal, which indicates ongoing data transmission, without a L0 input trigger from the TFC system,

(ii) FIFO overflow: there is a trigger but no data valid signal from the ARx cards.

The ADC FIFO error flag is set each time the FIFO overflows or underflows and as such indicates a malfunctioning of the FEM.

The occurrence of ADC FIFO errors was observed only once. Nine errors were observed when analyzing a long run taken in preparation for the TED run (run 58844) in October 2009. For this run NZS data was taken at a rate of 10 Hz and 123k events were collected. The ADC FIFO errors observed for this run are shown in Fig. 3.13.

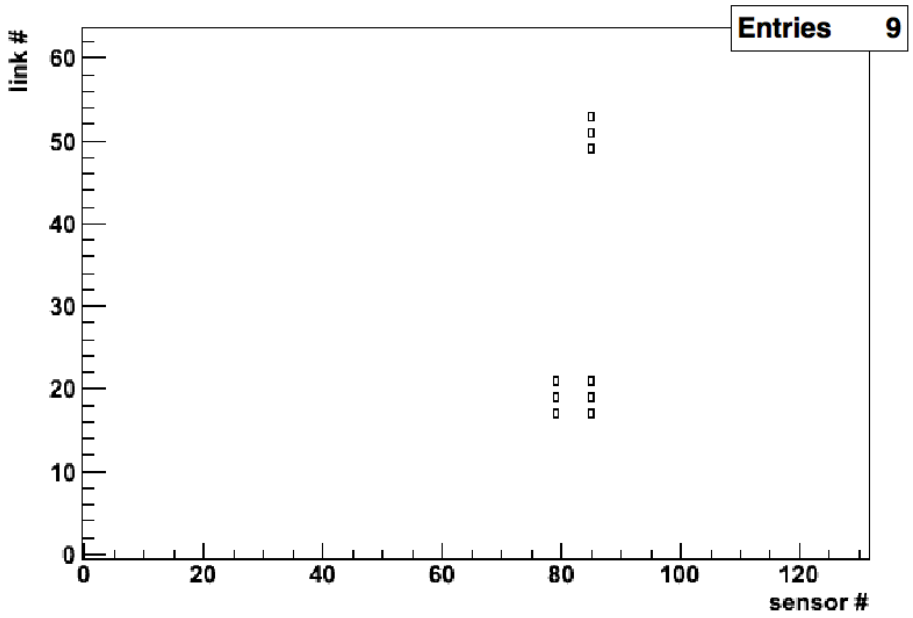

Figure 3.13: *ADC FIFO errors per sensor per A-Link.*

3.4.4 Channel error

The Channel error summarizes an overall error in the synchronization. It is a 4-bit number, one per PP_FPGA, corresponding to the Beetle number on a sensor (bit0 corresponds to Beetle0, bit1 to Beetle1, etc...) and flags if the ADC FIFO error, Pseudo header error flag, or PCN error is set.

3.4.5 IHeader error

To each Beetle chip corresponds an IHeader. This is an 8-bit number, labelled as $I_0 - I_5$ and $S_0 - S_1$ (see Sec. 3.1.3). In the raw data, the IHeader value is given by:

$$S_1\ S_0\ I_5\ I_4\ I_3\ I_2\ I_1\ I_0\ ,$$

where S_1 is the MSB and I_0 the LSB. To avoid mistakes due to a single event upset (SEU), the Beetle chip uses triple redundant logic. This consists of a logic bit that is represented by the majority of the output of three flip-flops. The static flip-flops use majority voting with a self-triggered correction mechanism. This means that if the output of the three flip-flops is not the same, the one that is different is changed to match the output of the other two. This is known as Error Detection and Correction (EDC) logic. If the 'ActiveEDC' bit (I_2) is set to 1 then the 'correction' logic is enabled. When something is being written to the chip (e.g. during configuration) this correction is disabled, due to the fact that unstable settings might induce bit flips that are not caused by an SEU.

IHeader errors appear in case the IHeader values of the 16 Beetle chips of one sensor are not equal. This occurs when an SEU is registered as the two MSBs of the IHeader correspond to the two LSBs of the SEU counter. Therefore, if an SEU is registered this counter is incremented and bits S_1 and S_0 of the IHeader are altered (see Sec. 3.1.3).

IHeader errors also appear if a pseudo header error is present, since in this case the value of the IHeader bits will be wrong. In Fig. 3.14 the number of IHeader errors per sensor per Beetle chip is depicted. In the current situation, IHeader errors deriving from the presence of pseudo header errors have been resolved.

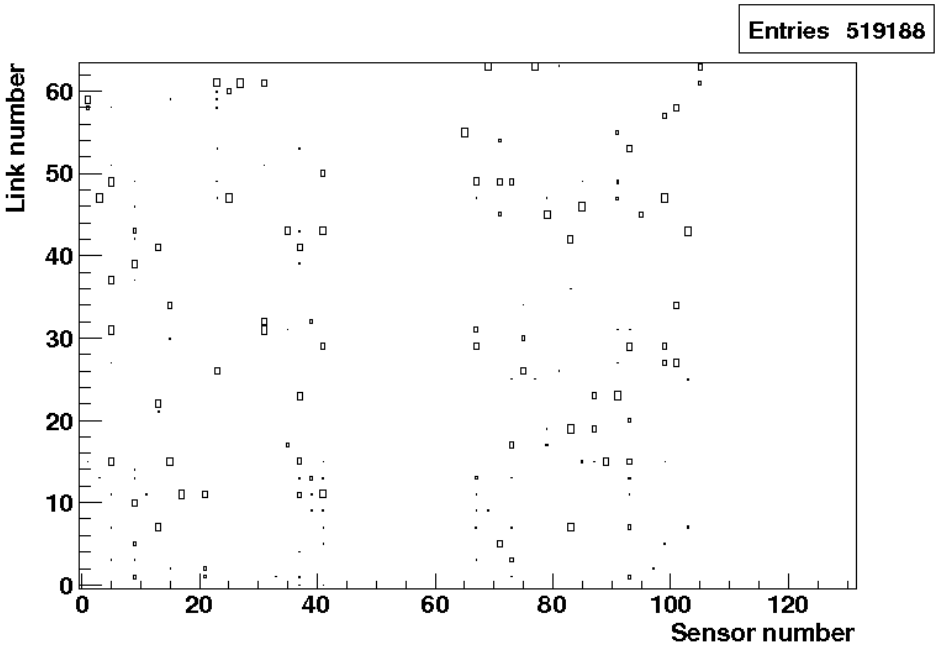

Figure 3.14: *IHeader errors per sensor per A-Link present at the beginning of the error analysis.*

3.4.6 Synchronicity between modules

The synchronicity between modules was checked to see if, for all modules, the PCN value is the same for a given event. To do this, the most frequent PCN value is determined among the FEM PCNs of all TELL1s. All the Beetle PCN values for a given event are then compared to this most frequent FEM PCN value, and if these values are not all identical an error flag is set.

The synchronicity studies were performed on the data run 62774, on approximately 11,000 events. All Beetle chips belonging to one module and all modules within one event were found to be synchronous.

3.5 Error monitoring

The error banks have to be monitored continuously due to changing beam conditions and parameters such as gain, pedestals, timing, etc. The analysis of the errors has been included both in the online monitoring and in the offline monitoring Graphical User Interface (GUI). The data that is processed in the online monitoring has passed the HLT. The HLT has an output rate of 5 kHz while the monitoring farm only analyses a fraction of this data. Each monitoring task has a different rate, depending on the task's complexity and processing speed. The VELO error monitoring runs on a dedicated task, which allows it to process data at a rate of 100 Hz. In this section a description of the plots that are generated is given. The plots shown in this section were produced from the data of run 58241, which consists of NZS data and contains approximately 80k events.

3.5.1 Online monitoring

The summary plots that have previously been described are generated and published per run in the online presenter. In addition, the statistics of occurrence of errors in all sensors are collected. In Figs. 3.15 and 3.16, two overview pages in the 'Error Banks' tab of the Velo Monitoring GUI are shown, for the arbitrary run 124278.

In Fig. 3.15 overview plots showing pseudo header errors, PCN errors, ADC FIFO errors and IHeader errors are shown. Each of these errors is identified per sensor and per A-Link. From the figure it can be seen that three TELL1s were not yet properly tuned for this particular run. In particular, TELL1 85 has errors on 4 consecutive A-Links, due to the fact that there is a dead Beetle chip for this sensor.

In Fig. 3.16 the overview page of the error banks monitoring in the GUI is shown. This page contains plots showing the number of errors present per sensor, the rate of errors per sensor, and which errors are present for each sensor.

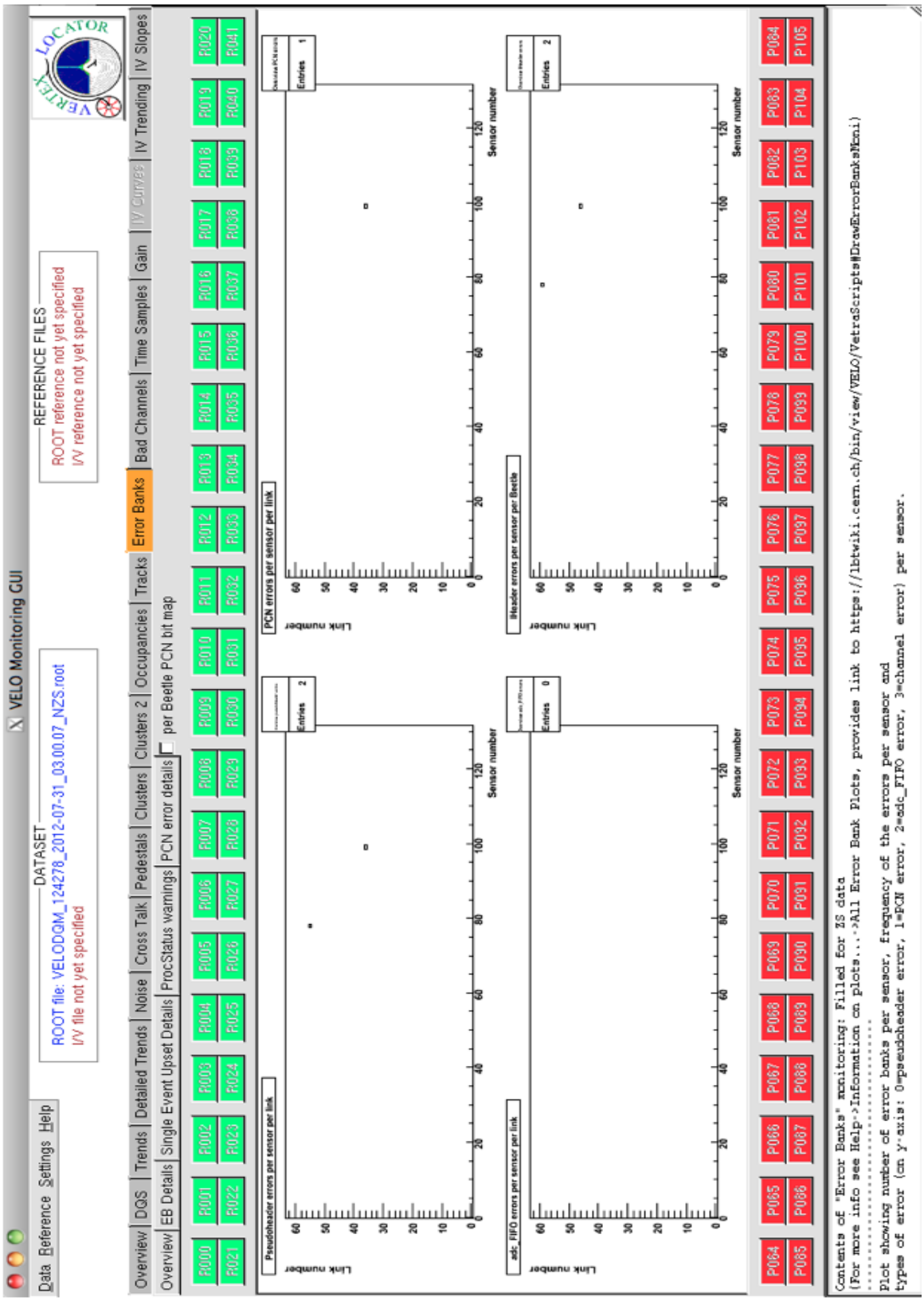

Figure 3.15: *'EB Details' page of the Error Banks tab in the Velo Monitoring GUI. Summary plots of the number of pseudoheader (top-left), PCN (top-right), ADC FIFO (bottom-left), and IHeader (bottom-right) errors which are present per sensor per link are shown.*

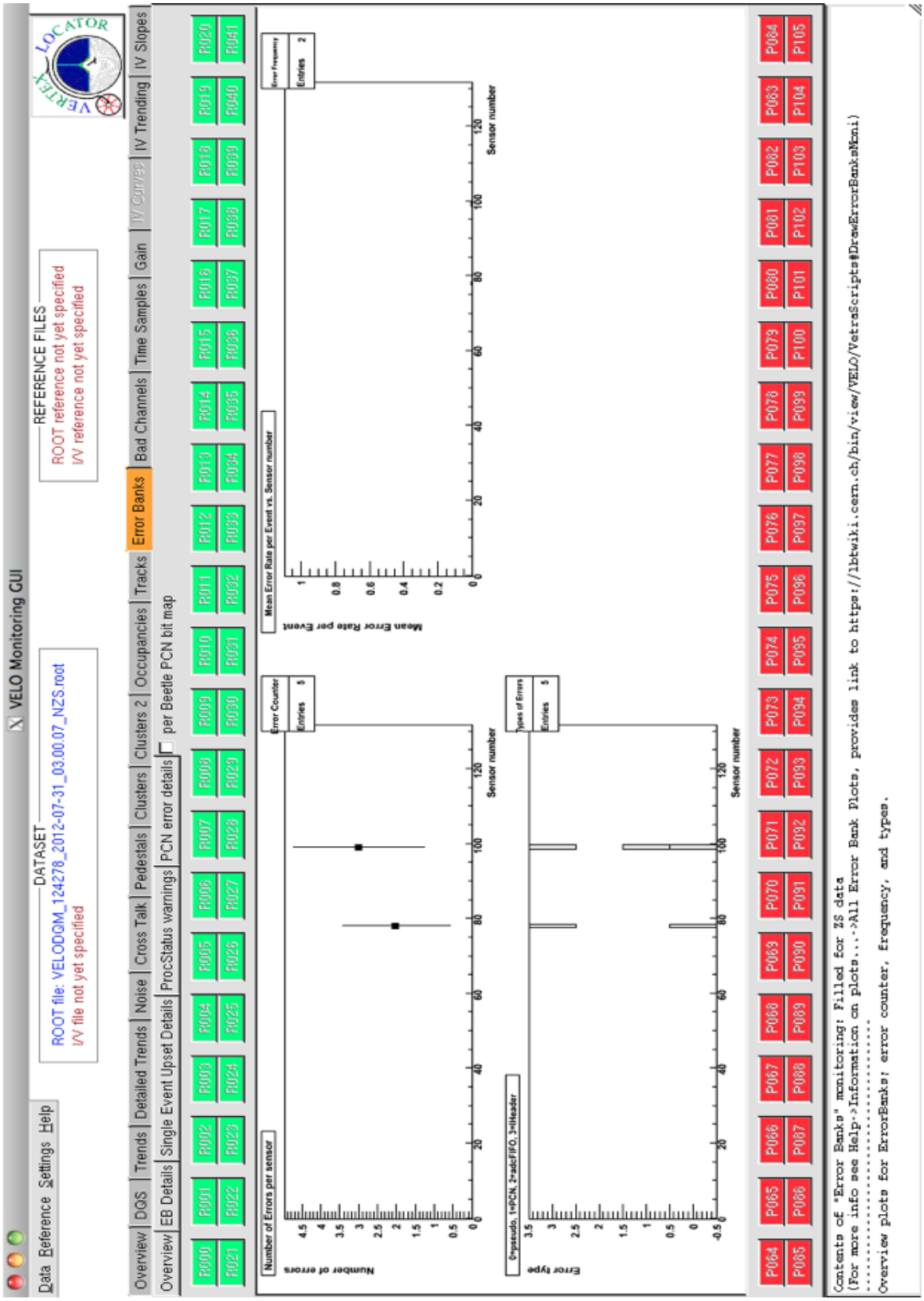

Figure 3.16: *'Overview' page of the Error Banks tab in the Velo Monitoring GUI. The number of errors per sensor (top-left), rate of errors per sensor (top-right), and which errors are present per sensor (bottom-left) are shown.*

In the bottom-left plot of this figure, a more condensed overview of which errors are present for each sensor can be seen. Along the x-axis the sensor number is displayed, along the y-axis the different types of errors. The following y-values correspond to the following errors:
· '0' = Pseudo header errors
· '1' = PCN errors
· '2' = ADC FIFO errors
· '3' = Iheader errors

The distributions of the number of errors present per sensor and of the fraction of the error banks for each sensor are also included in the online monitoring, and are shown on the top in Fig. 3.16. The fraction of errors is useful to make different data sets directly comparable.

Lastly, the number of errors that is present per sensor in the last N events of the run is published in the online monitoring. N is currently set to 10000 but this value can be modified in the code for the error monitoring.

The final plot that is included in the online monitoring is used to check that there is synchronicity within modules, meaning that for one TELL1 the PCN values of all Beetle chips are equal to the value of the FEM PCN. If the PCN of one or more Beetle chips of one module is different from the FEM PCN, then the number of times the TELL1 is asynchronous is counted for each TELL1 and the wrong PCN value causing this asynchronicity is indicated. The number of sensors that have errors is also published in the online monitoring.

3.5.2 Offline monitoring

While the online monitoring runs automatically during data taking, the offline monitoring can be run on any given file. This is useful because in the online monitoring only a small subset of the data is analysed. When data is taken it is saved by run; these files are stored and can be inspected with the monitoring GUI (see figures 3.15 and 3.16). The offline monitoring allows for the selection of the number of events on which to run on or of the types of data banks to be analysed.

3.5.3 Summary and current status

Error banks are present in approximately 0.1% of the events. The presence of an error bank, however, does not imply that a specific event must be excluded from the final collected data sample. In September 2009, the frequency of the errors was on average 10^{-3}, with errors appearing in all TELL1s[3]. Most errors were found to be caused by an incorrect tuning of the pseudoheader thresholds, which leads to the presence of pseudoheader errors and, consequently, PCN and IHeader errors.

Between 2010 and 2011, a number of A-Links, Beetle chips and TELL1s were found to be defective. Problematic TELL1 modules were substituted and specific Beetle chips and A-Links were masked, in order to exclude the data from these problematic components. A list of known problems is given in Table 3.2. Currently, error banks are observed at a rate of approximately 10^{-3} only on specific TELL1s (normally only one or two sensors have errors in the same run). Since error banks are usually only present on specific A-Links or Beetle chips, events containing error banks are still included in data taking. Based on the low rate at which error banks are observed it can be concluded that the quality of the data taking process at LHCb is not affected.

[3]For various sensors the rate of error banks was as high as 10^{-2}.

Affected component	Description
Sensors 0+10	Links with stuck bits
Sensor 67, link 31 (ch.992−1023)	Very noisy(raw noise), masked
Sensor 82, link 49 (ch.1568−1599)	Zero noise, masked
Sensor 82, link 51 (ch.1632−1663)	Half-gain link
Sensor 85, links 28−31	Dead Beetle chip
Sensor 85, links 52−55	PCN and channel errors
Sensor 97, links 48−51	Beetle 12 gets out of sync.
Sensor 99, link 20 (ch.640−671)	Half-gain link
Sensor 104, link 19 (ch.608−639)	Desync errors in SEU monitoring; half-gain link

Table 3.2: *List of known problematic components and related issues.*

Chapter 4

Signal Distributions of Headers and Channels

In this chapter the analysis performed to study the origin of pseudo header errors is described. For each A-Link, the analogue header values were studied in combination with data of the preceding and subsequent analogue channel. In order to analyse the ADC values of headers and strip-channels, Non-Zero Suppressed (NZS) and Time Aligned Event (TAE) data are required. In Sec. 4.1 the event selection is described. A possible source of pseudo header errors is investigated in Sec. 4.2 through the analysis of header bit 3 and header bit 0. In Sec. 4.3, a similar analysis is performed on ϕ-sensors only, and the differences between signals on adjacent and non-adjacent strips are studied. An effect known as *JC-Effect* is described, calculated and corrected for. Finally, in Sec. 4.4 a summary of the results is given and conclusions are drawn.

4.1 Selection of Events

Data is acquired in different ways in the VELO detector, one of which is known as Round Robin mode. In this mode, only 1 of every 1000 events contains an NZS bank for a specific TELL1. Run 81811, which was used for this analysis, was taken in Round Robin mode. Of the approximately 5 million events in the run 128711 NZS banks are available. The analysis presented here was performed on these banks.

It was observed that out of these $\simeq$ 129k events only 20 events contained both an NZS bank and an error bank for the same TELL1, which is insufficient for an analysis. To overcome this lack of statistics, all events containing an NZS bank were selected and stored in a new raw data file, such that all 129k events are used for further study.

After all events were analysed, it was observed that the occurrence of a pseudo header error is always caused by a faulty value of either header bit 0 or header bit 3. In the following sections the behaviour of these bits will be described separately.

4.2 Analysis of header bit ADC values

Each A-Link contains 36 values, 4 header bits (numbered $0-3$) and 32 strip channel values ($0-31$). In the following a possible correlation of the ADC value of the header bit and an adjacent channel is studied. This is most straight-forward for header bit 3 since it is followed by the first channel ADC in the same event; while it's more involved for header bit 0, which is preceded by the last channel ADC of the previous event.

4.2.1 Analysis of Header bit 3

The baseline ADC value of a header or analogue channel is typically around 512 ADC counts. For the header bits, a digital value of '0' is assigned when the ADC value is higher than the *high pseudo threshold*, and '1' when the ADC value is lower than the *low pseudo threshold*. The determination of these pseudo thresholds was described in the previous chapter and illustrated in Fig. 3.10. The separation is typically 40 ADC counts or less.

In principle, headers and analogue channels should be independent. However, a correlation is observed between the ADC value of header bit 3 and that of strip-channel 0, in particular when the strip ADC has extreme values.

To study this effect the analogue channels are classified in such a way that the signal on strip-channel 0 (channel number 4) is defined *high* when the ADC corresponds to more than 3 times 1 MIP (around 50 ADC counts). Alternatively, it is classified as *negative* if it is significantly lower than the baseline, typically around 400 ADC counts. In Fig. 4.1 these two cases are depicted.

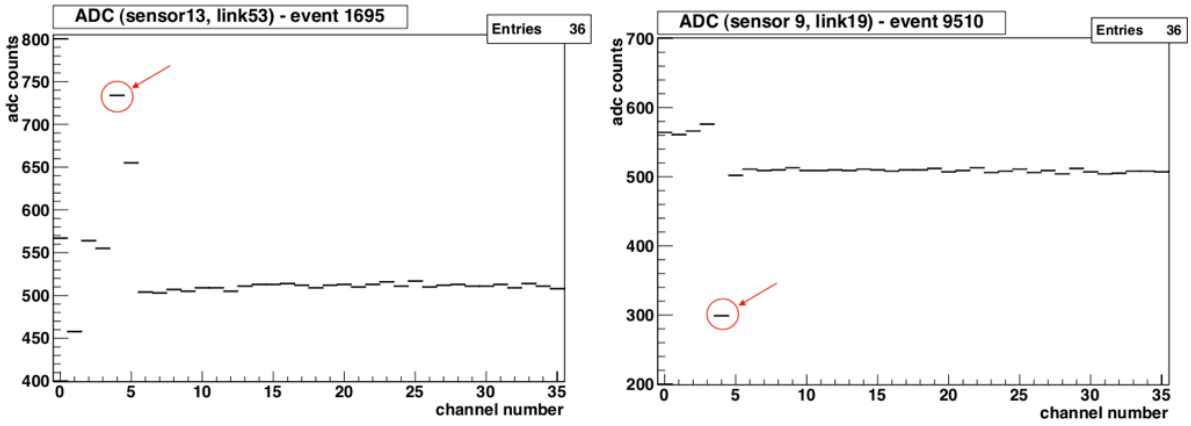

Figure 4.1: *ADC values for headers and strip-channels for two cases. Channel numbers* $0-3$ *indicate the header bits and* $4-36$ *indicate the 32 analogue strip-channels of an A-Link. The high (left) and negative (right) ADC values of strip-channel 0 (channel number 4) are indicated with an arrow.*

It must be taken into consideration that the value of the baseline varies between different A-Links, Beetle chips and TELL1s. In order to analyse the data of the whole VELO, the pedestal for each A-Link was calculated and then subtracted from the ADC values of that A-Link.

The procedure to calculate the pedestal was as follows. For each A-Link, the average of four strip-channels (8, 9, 10, 11) was calculated over all events. The strip-channels used to calculate this average are chosen arbitrarily, excluding channels at the 'edges' of the series to avoid possible cross-talk with neighbouring A-Links. The resulting values are stored in a histogram and the most probable value of this average is taken as the pedestal. This average value is only valid in first approximation, but is acceptable for the purpose of this study. By choosing the most probable value of this distribution high and negative ADC values are neglected, thus preventing this measurement from being biased by the Landau tail of the distribution. The pedestal values for all A-Links of one module are shown in Fig. 4.2.

For each TELL1, the maximal difference between the pedestals of the 64 A-Links was calculated. This maximal difference between pedestal values for all TELL1s is shown in Fig. 4.3.

From the figure it can be seen that the pedestal values of different A-Links of one TELL1 can differ by as much as 48 ADC counts. As was described in the previous chapter, until March 2011 the header thresholds were set per TELL1. The distance (in ADC counts) between the

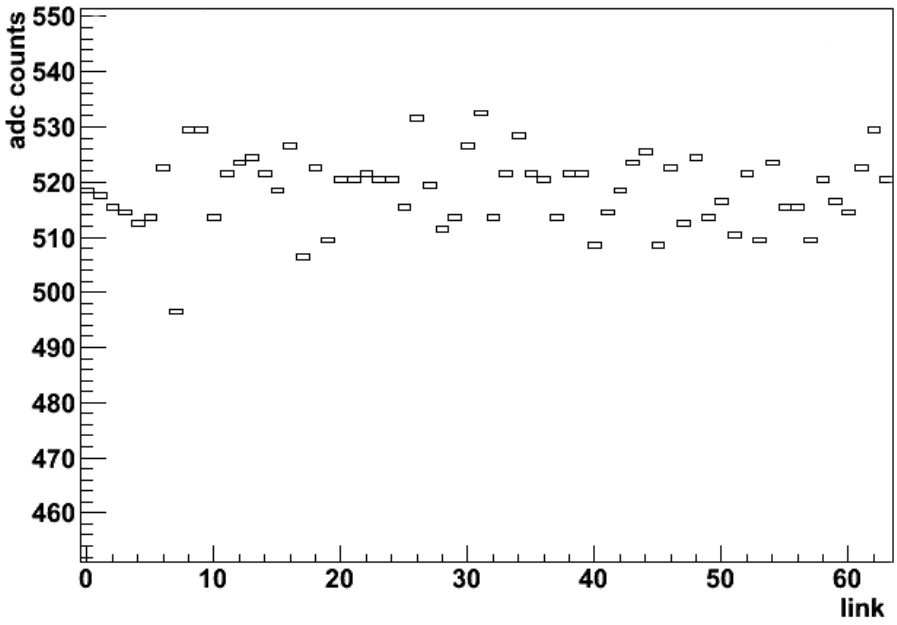

Figure 4.2: *Pedestal value per A-Link calculated over all events of sensor 4.*

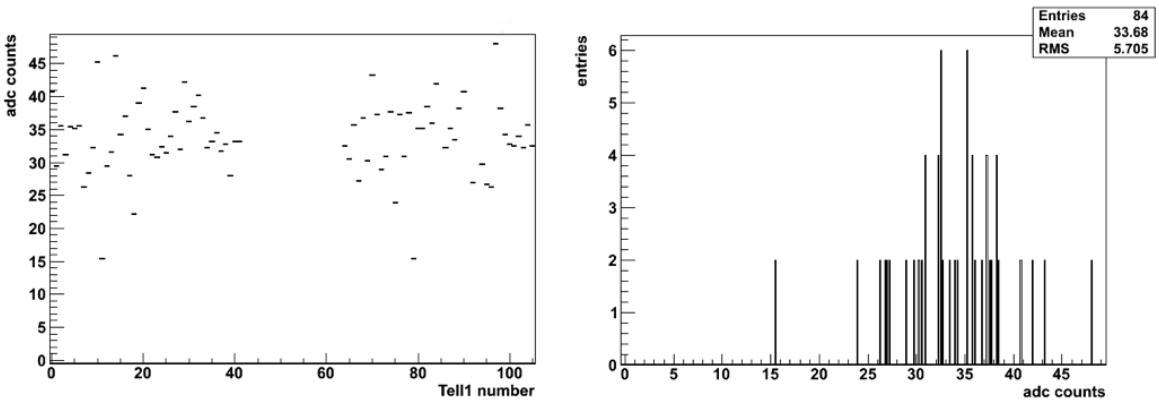

Figure 4.3: *Left: Maximal difference between A-Link pedestals vs TELL1. Right: Distribution of maximal difference between A-Link pedestals per TELL1.*

high and low header thresholds is on average approximately 40 ADC counts, but can also be less (see Fig. 3.10 chapter 3). This means that there are cases in which the offset between different A-Links is larger than the distance between the thresholds. In these cases it is impossible to distinguish between 'real' header errors and errors due to the limited granularity of the threshold settings.

This problem can be solved by setting the header thresholds per A-Link instead of per TELL1. This feature has been included in the latest version of the firmware[1] and the new set of header thresholds has been implemented in the database.

To be able to combine the information of each strip-channel 0 of all A-Links of all TELL1s, the pedestal value is subtracted from the ADC values of all strip-channels. In Fig. 4.4 the correlation between the ADC value of header bit 3 and strip-channel 0 is shown. It can be seen

[1]version 4.2 of the VELO TELL1 firmware contains this feature.

from this figure that header bit 3 and strip-channel 0 are inversely correlated; a high value of strip-channel 0 correlates to a value of header bit 3 that is lower than the average. To a negative value of strip-channel 0 corresponds a value of header bit 3 that tends to be somewhat higher than the average. The variation of the ADC value of header bit 3 from the average of all header bit 3 values has been calculated in first approximation to be about −5% of the (ADC−pedestal) value of strip-channel 0, for both high and low headers.

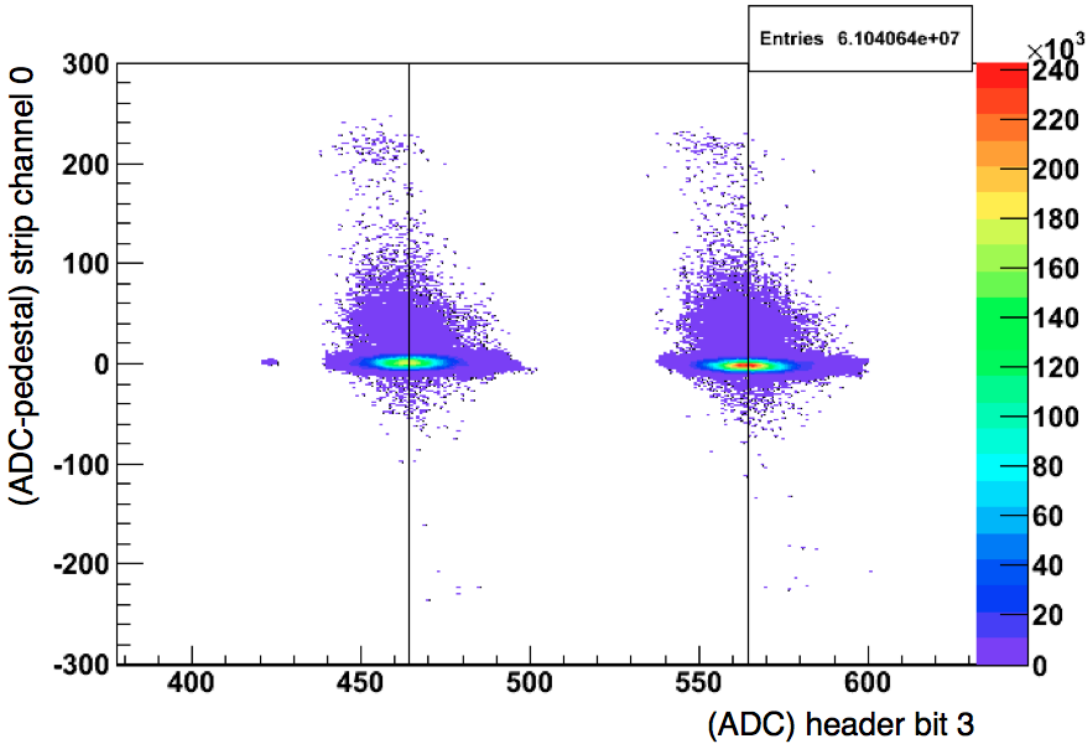

Figure 4.4: *Correlation between ADC values of strip-channel 0 and header bit 3.*

Figure 4.5 shows a zoom-in along the vertical (ADC−pedestal) axis. In this plot it can be seen that the central value of strip-channel 0 appears to be shifted comparing the cases where the header bit 3 was low to that where it was high. The most probable value should be 0, corresponding to the baseline. However, when the value of header bit 3 is low, around 464 ADC counts, the peak is at 1.6. Similarly, when the value of header bit 3 is high (around 564 ADC counts), the ADC distribution of strip-channel 0 is pulled down to about −2.

With the position of the center of each header distribution, the following ratio was calculated:

$$\frac{(ADC - pedestal)_{High} - (ADC - pedestal)_{Low}}{(Header)_{High} - (Header)_{Low}} \tag{4.1}$$

where the indices 'High' and 'Low' denote the center of the most dense region of the high and low header distributions, respectively. The ratio between the centers of these distributions was found to have a value of approximately −4%. The fact that this ratio is comparable to the cross-talk between strip-channel and header hints to the presence of a similar effect as the cause of this shift. The direction of the shift between the center of the two distributions is also in agreement with the trend that was observed: to lower header values are associated higher (ADC−pedestal) values of the 'neighbouring' analogue strip-channel. As in the case of channel-header cross-talk, this shift can be attributed to a forward cross-talk between headers and analogue strip-channels.

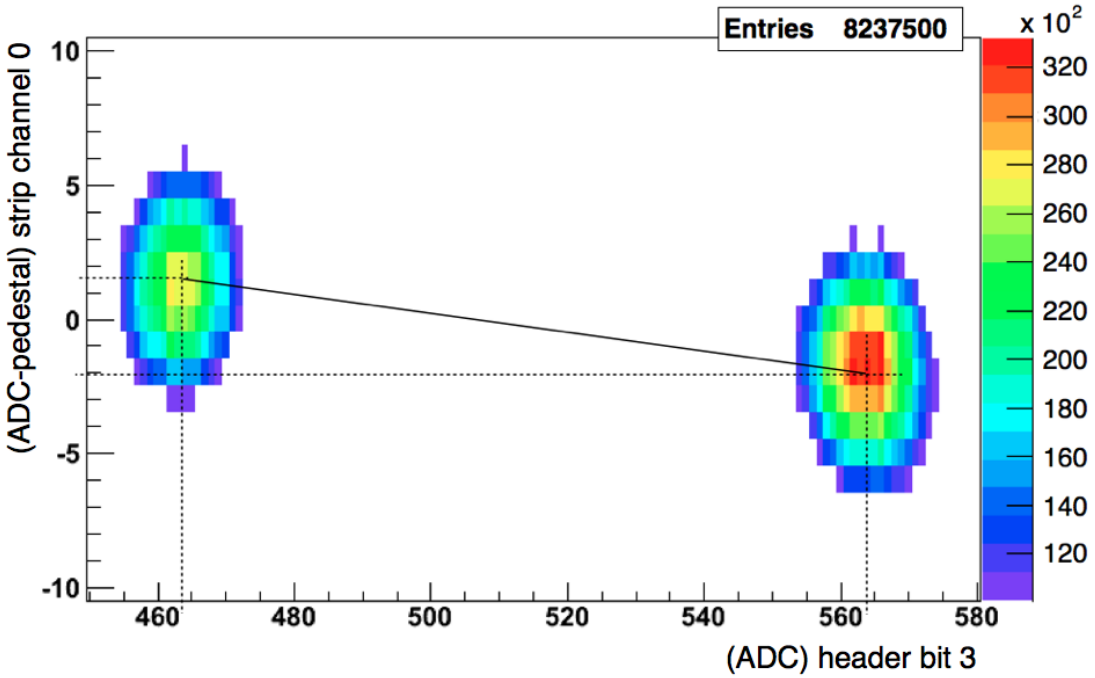

Figure 4.5: *Zoom of the most populated region of the high and low header distributions. The distributions are shown for the case of header bit 3 vs channel 0.*

4.2.2 Analysis of header bit 0

In order to study a correlation between strip-channel 31 and header bit 0, events that are consecutive in time are required. For this purpose Time Aligned Events (TAE) are used. TAE data is normally used to determine the optimal sampling time of a pulse shape (which is sampled once per bunch-crossing, i.e. every 25 ns). These sets contain five triggers, corresponding to five events that are consecutive in time. In this analysis run 70337, containing NZS data, was used. This run has optimal timing and contains 19,964 events per time sample, corresponding to a total of approximately 100,000 events.[2]

For each A-Link, the ADC values of the 4 header bits and 32 analogue strip-channels of each time sample are studied. In Fig. 4.6 an example of five consecutive events is shown. Along the x-axis of this plot it can be seen that the 4 header bits and 32 analogue channels of each time sample are present ($0-3$ are the header bits and $4-35$ are the analogue strip-channels of time sample 0, $36-39$ are the header bits and $40-71$ are the analogue strip-channels of time sample 1, etc...).

With these data, the correlation between the ADC value of strip-channel 31 of time sample N and that of header bit 0 of time sample $N+1$ can be studied.

In Fig. 4.7 the ADC values of the headers and analogue strip-channels of one A-Link are shown. Two different cases are shown: in the first there is a high hit ((ADC−pedestal)>150) on strip-channel 31 of time sample 1, while in the second there is a negative hit on this channel also in time sample 1((ADC−pedestal)<−100). In Fig. 4.8 the correlation between the ADC value of strip-channel 31 and that of header bit 0 is shown, for all TELL1s.

When the ADC value of strip-channel 31 is negative, the value of header bit 0 is higher than its average value for that link. Similarly, when the value of strip-channel 31 is high, the value

[2]Event 583 was discarded because it was found to be corrupt. Data of link 31 of sensor 67 and link 38 of sensor 70 were also excluded.

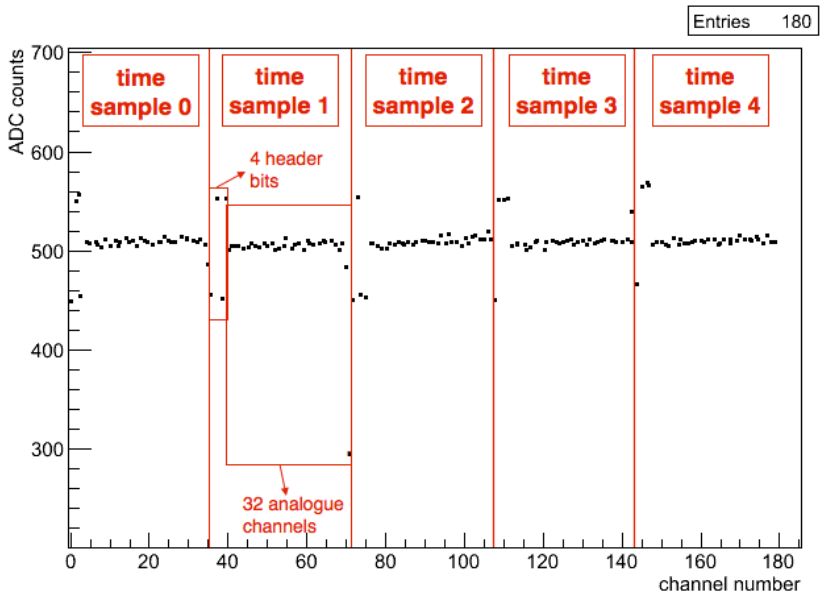

Figure 4.6: *Definition of channel numbers in case of 5 consecutive events. The channel numbers include both header bits and analogue strip-channels.*

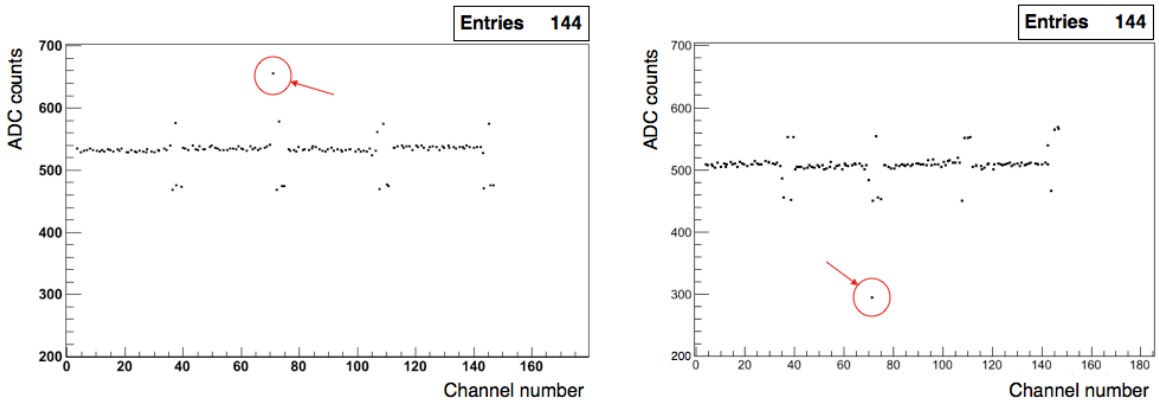

Figure 4.7: *ADC values of headers and analogue channels for one event for one A-Link for which the (ADC−pedestal) of strip-channel 31 is: > 150 ADC counts (left), and <−100 ADC counts (right).*

of header bit 0 is lower than its average value for that link. The slope of this correlation is such that the amplitude of header bit 0 changes by about −4% of the deviation from the average in strip-channel 31. This value is comparable to that found for header bit 3 and strip-channel 0.

Similarly, cross-talk between header bit 0 and strip-channel 31 is observed. The MPV of the ADC distribution of high headers is lower than that associated to the low header distribution, as can be seen in Fig. 4.9, a zoom-in of the vertical scale of Fig. 4.8. The ratio between the centers of the low and high header distribution is equal to approximately −6%, showing that there is backward cross-talk from header to channel.

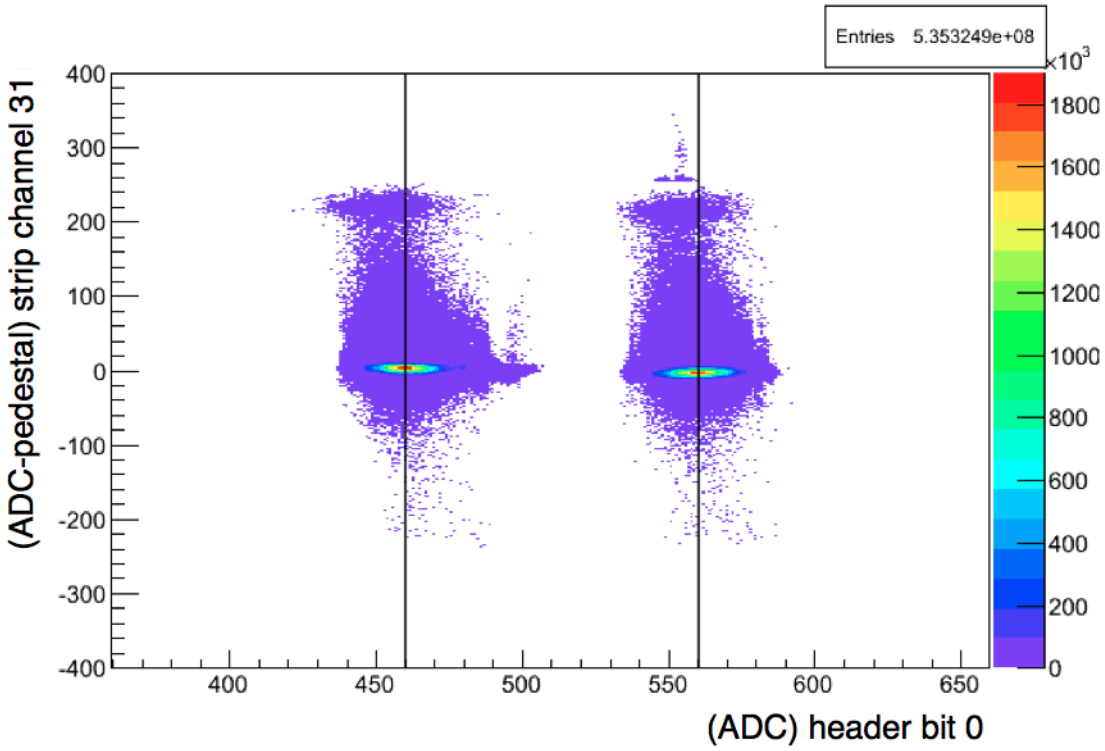

Figure 4.8: *Scatter plot of the ADC values of strip-channel 31 and header bit 0.*

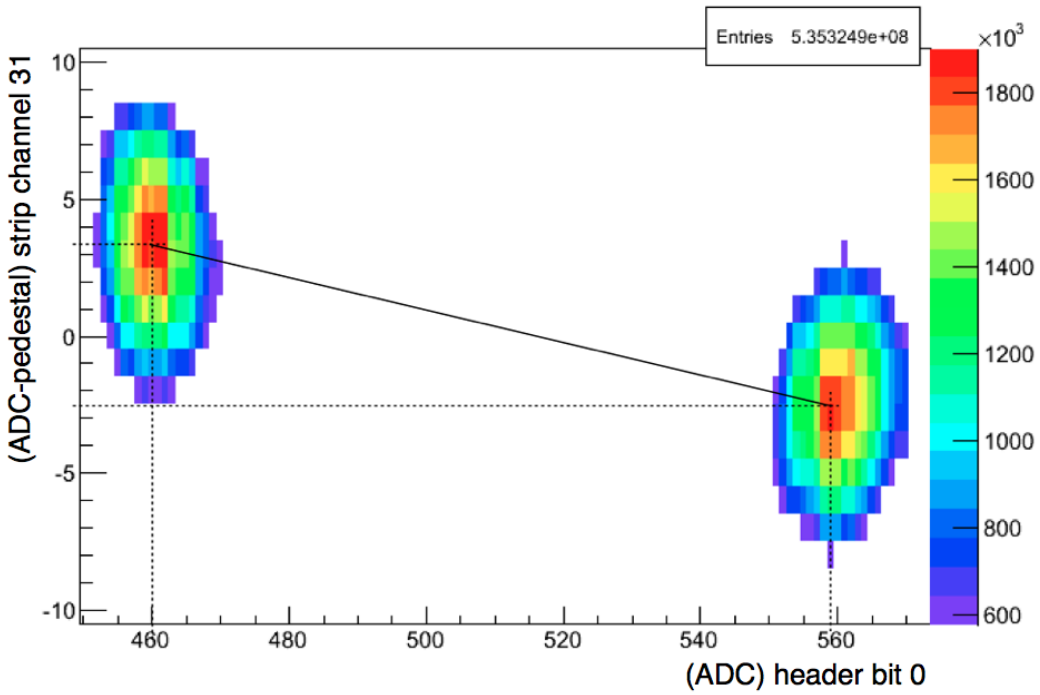

Figure 4.9: *Zoom of the most populated region of the high and low header distributions. The distributions are shown for the case of header bit 0 vs strip-channel 31.*

4.3 Study of ϕ sensors

To further investigate the correlation between the values of header bit 3 and strip-channel 0, ϕ-sensors were analysed in more detail. The reason for choosing the ϕ-sensors is that on these sensors the strips are separated in inner and outer regions. As a consequence, for these sensors consecutive readout channels do not always correspond to adjacent strips on the sensor. This allows to see whether the phenomenon described in the previous section has its origin in the sensor or in the DAQ chain, for example, in channel-to-channel cross-talk.

Neighbouring channels were studied to investigate if a high ADC value of one channel has an effect on the ADC values of the previous and next channels. If, for example, strip-channel 2 is selected, the ADC distributions of strip-channel 1 and strip-channel 3 are studied for all events for which the (ADC−pedestal) value of strip-channel 2 is >100 ADC counts. There is a pattern of correspondence between strip-channels and strips that repeats itself every 12 channels. This pattern is shown in Table 4.1. Each ϕ-sensor has 2048 channels, so in order to consider all channels corresponding to the same strip class the channel numbers are classified as:

- strip-channel numbers for which *strip-channel number modulo 12 = 0* are indicated as 'type 0 channels'
- strip-channel numbers for which *strip-channel number modulo 12 = 1* are indicated as 'type 1 channels'
- strip-channel numbers for which *strip-channel number modulo 12 = 2* are indicated as 'type 2 channels'
- etc...

channel type	strip	previous	next
type 0 channel	inner strip	non-adjacent	non-adjacent
type 1 channel	outer strip	—	—
type 2 channel	outer strip	—	—
type 3 channel	outer strip	—	—
type 4 channel	inner strip	non-adjacent	adjacent
type 5 channel	inner strip	—	—
type 6 channel	outer strip	—	—
type 7 channel	outer strip	adjacent	adjacent
type 8 channel	outer strip	—	—
type 9 channel	inner strip	—	—
type 10 channel	outer strip	—	—
type 11 channel	outer strip	adjacent	non-adjacent

Table 4.1: *Correspondence between channel numbers and region on the sensor in which the strip is located. For the four channels studied it is indicated whether the previous and next readout channels correspond to adjacent or non-adjacent strips on the sensor.*

In Figures 4.10 (a)−(d), the ADC distributions of neighbouring channels are shown for four different situations. The channels for which the (ADC−pedestal) value is required to be >100 ADC counts are in turn type 0 (Fig. 4.10(a)), type 4 (Fig. 4.10(b)), type 7 (Fig. 4.10(c)) and type 11 (Fig. 4.10 (d)). In this figure, the distribution on the left contains pedestal subtracted ADC values that pass the cut (ADC−pedestal)>100 of the channel under examination, while on the right the ADC distributions of the two adjacent channels are shown.

Channels of type 0 correspond to inner strips. The previous (type 11) and next (type 1) channels correspond to outer strips, such that these readout channels are non-adjacent strip

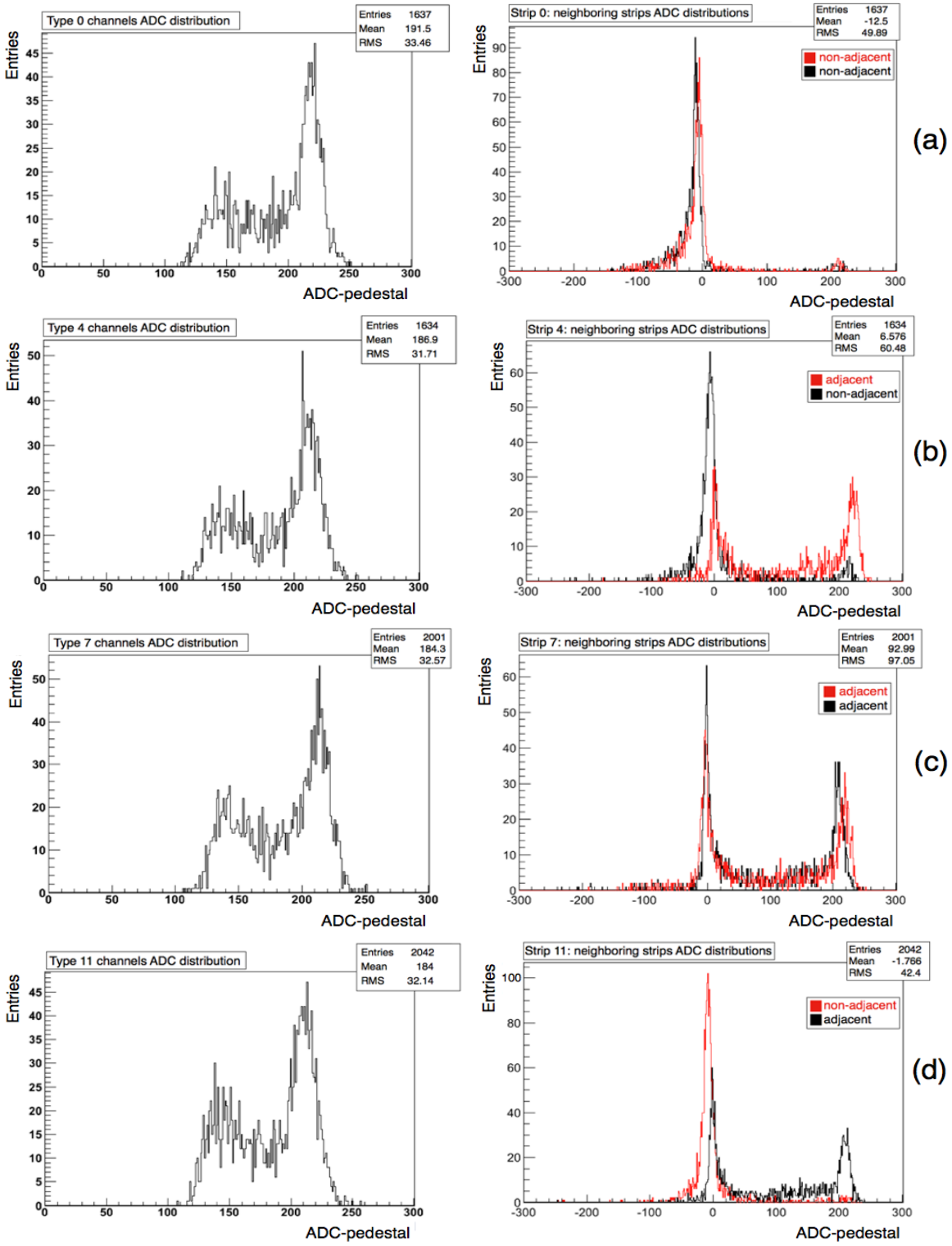

Figure 4.10: *Left: ADC distribution of (a) type 0, (b) type 4, (c) type 7 and (d) type 11 channels. Only events with (ADC−pedestal)>100 are selected. Right: in black the ADC distribution of previous channels, in red the ADC distribution of next channels. If the strip is adjacent or not to the studied one is indicated in the plots.*

positions on the sensor. Type 4 channels correspond to inner strips. In this case, the previous channel (type 3) corresponds to a non-adjacent outer strip, while the next channel (type 5) corresponds to an adjacent inner strip. In the case of type 7 channels, both previous and next channels (type 6 and 8) correspond to adjacent outer strips. Finally, the last case investigated is that of type 11 channels. In this case, the previous channel (type 10) corresponds to an adjacent outer strip, while the next channel (type 0) corresponds to a non-adjacent inner strip. This is summarized in Table 4.1.

The value of the peak of the ADC distributions for different channels is given in Table 4.2. It is also indicated in the table whether the neighbouring channels correspond to adjacent or non-adjacent strips.

channel	position of peak	adjacent strip	ref. to distribution
0 (previous)	-12	No	Fig. 4.10(a), black curve
0 (next)	-5	No	Fig. 4.10(a), red curve
4 (previous)	-8	No	Fig. 4.10(b), black curve
4 (next)	0	Yes	Fig. 4.10(b), red curve
7 (previous)	-2	Yes	Fig. 4.10(c), black curve
7 (next)	-2	Yes	Fig. 4.10(c), red curve
11 (previous)	0	Yes	Fig. 4.10(d), black curve
11 (next)	-10	No	Fig. 4.10(d), red curve

Table 4.2: *Characteristics of the ADC distributions for the studied channels with respect to the corresponding strip position on the ϕ-sensor.*

There are two separate features that can be seen in the figures. The first is due to charge sharing between strips on the sensor, leading to a peak at high ADC values, and the second is due to a correlation between readout channels, leading to shifted pedestals.

The peak around 200 ADC counts in the distributions of channels corresponding to adjacent strips. This peak is caused by two effects:

- There is charge sharing between adjacent strips, which varies between 0% and 100%. As a consequence, it is possible that strips adjacent to a strip with a hit also have a high ADC value, causing a long tail of high ADC values in the distribution.

- Since the Beetle chip saturates at a value of approximately $200-250$ ADC counts above the baseline, all higher ADC values are concentrated at this limit, which causes a second peak in the distribution. The width of this peak is due to the fact that the saturation point differs per Beetle chip.

The main feature of interest is a shift in the pedestal values. In the plots it can be seen for the cases where the neighbouring channels correspond to non-adjacent strips. When the neighbouring channels correspond to non-adjacent strips, and the (ADC$-$pedestal) value of one channel is >100, then for the same event the ADC values of neighbouring channels are shifted towards lower values. This effect is visible on the righthand of Fig. 4.10 for: (a) both distributions, (b) distribution in black (previous channel) and (d) red distribution (next channel).

The fact that the same negative shift is not visible in adjacent strips implies that charge sharing compensates for it, and does not necessarily indicate that this negative shift does not exist. However, in this case the effect cannot easily be separated from the charge sharing on adjacent strips on the sensor.

For each of the 4 selected strip types, the correlation between the pedestal subtracted ADC value and the negative shift of 'non-adjacent' previous and next pedestals was measured. A

typical correlation of -4% to -6% was observed, which compares to the correlation of strip-channel 0 (31) and header bit 3 (0). The fact that it was observed in 'non-adjacent' strips confirms that, as is the case for strip-channels and headers, this effect has its origin in the readout chain and not in the sensor.

4.3.1 Calculation and Correction of the 'JC'-Effect

A possible cause for the shift in the peak of the ADC distributions is an effect known as the JC-effect (named after the person who discovered it), which occurs when a large signal causes all pedestals of an entire Beetle chip to shift downwards. The amount by which the pedestal of a Beetle chip shifts down is proportional to the amount of charge deposited in the sensor: the integral over the ADC values of channels with large signals and that of all other channels have to compensate each other.

An event for which the shift in pedestal is clearly visible is shown in Fig. 4.11. For this event, the JC-effect can be observed on Beetle chip 4, so for analogue channels 512−639. On this Beetle chip there are 37 channels with high ADC values ((ADC−pedestal)>150) and 7 channels with signals of nominal size. The fact that a large number of channels have an ADC higher than 150 indicates the presence of a large energy deposition in the sensor, which can be caused for example by delta rays. These signals are concentrated around 200 ADC counts due to the saturation of the Beetle chip. The presence of these signals causes the pedestals of the remaining 86 channels to be shifted downwards by about 40 ADC counts. The magnitude of this shift corresponds to the magnitude of the JC effect.

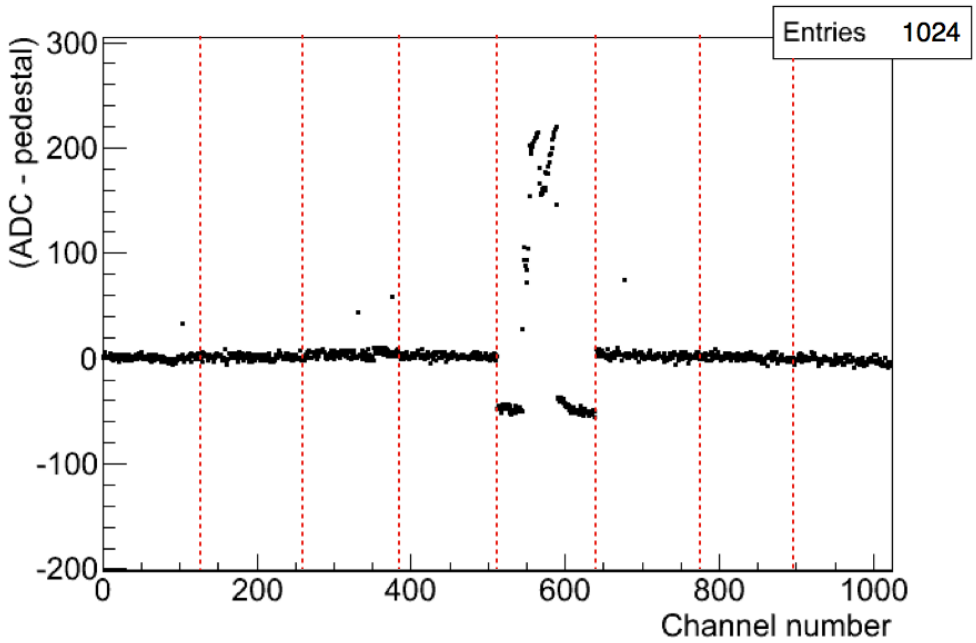

Figure 4.11: *Event for which the downward shift in ADC values due to the JC-effect is visible for the channels of Beetle chip 4. Channels are shown for Beetle chips 0−7, separated by red lines.*

The JC-effect is studied here by looking at the correlation between the signals of previous and previous-to-previous channels. The reason for studying these channels is that, if there is a high signal on one channel, the shift that might be observed in the previous channel can be caused

not only by the JC-effect, but also by other effects such as channel cross-talk. If however, for the same event, the value of the previous-to-previous channel is also shifted towards lower values, this will most likely be caused only by the JC-effect. The correlation between the (ADC−pedestal) value of the previous channel and that of the previous-to-previous channel, which are adjacent on the sensor, when selecting a high signal (>150) on a channel of type 4, is shown in Fig. 4.12. This means that two outer strip channels are tested for a high signal on an inner strip channel. Therefore, correlations as seen in the previous section are excluded.

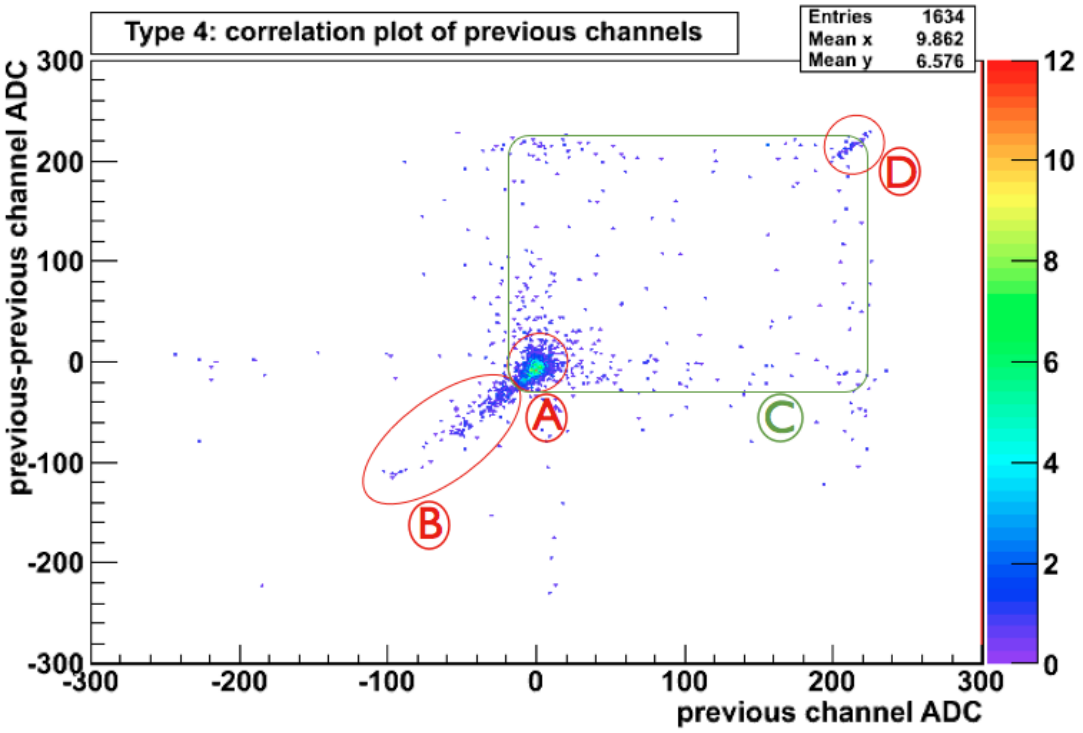

Figure 4.12: *(ADC−pedestal) distributions of previous channel vs. previous-to-previous channel, for type 4 channels. Only type 4 channels for which (ADC−pedestal)>150 are selected.*

In Fig. 4.12 four different regions are indicated:

region A Most of the data is concentrated around [0,0]. This shows that there is no signal on most of the strips. The spread in the data is due to noise.

region B Here the JC-effect can be seen qualitatively, where the ADC values of both the previous and previous-to-previous channel are negative with equal values (i.e. 100% correlated).

region C Hits on the strips can be seen in this region. Since the two strips are adjacent there is the possibility that a hit on one of the strips will leave a charge deposition on the other.

region D There is a concentration of data which is due to the fact that the Beetle chip saturates, so all signals with an ADC value higher than the saturation point will concentrate around one value (the spread is due to the fact that different Beetle chips have different saturation points).

In order to understand to what extent the shift in peak position towards lower ADC values is influenced by the JC-effect, this effect was investigated in a test pulse calibration run [54]. A known amount of charge was injected to a few channels of a Beetle chip. As a consequence, a shift in the baseline towards lower values was observed. The baseline shift in response to different magnitudes of injected charge is shown in Fig. 4.13.

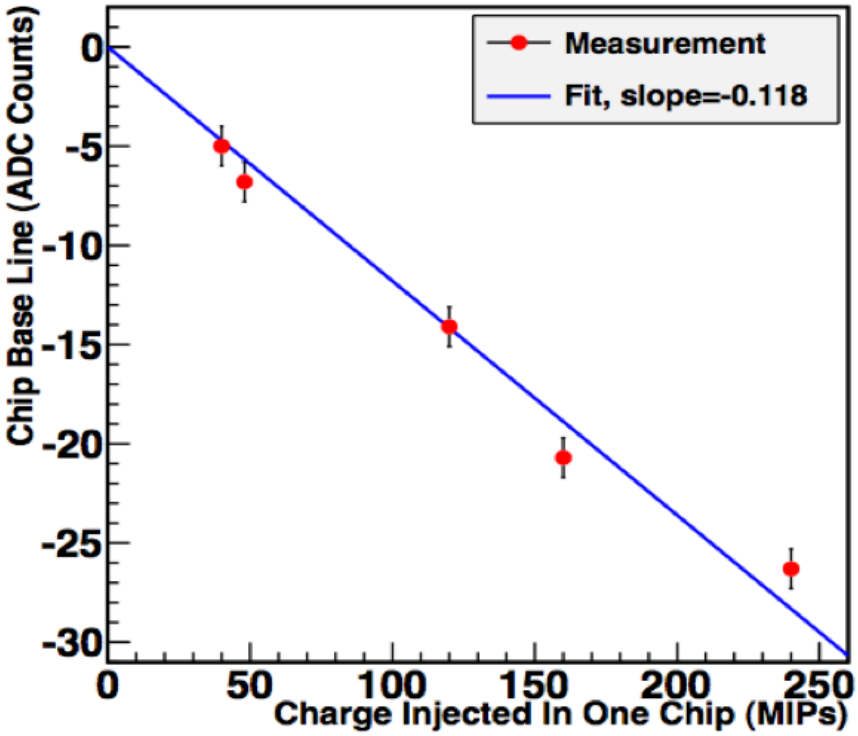

Figure 4.13: *Shift in baseline of one Beetle chip as a function of the injected charge (plot taken from [54]).*

For a total injected charge of 50 MIPs in one chip, corresponding to about 7000 ADC counts, this baseline shift is equal to -5 ADC counts. The shift in the baseline linearly increases to approximately -30 ADC counts for an injected charge of 250 MIPs. The JC-effect was 'quantified' per Beetle chip, using all events for which at least one channel has an (ADC$-$pedestal) value >200. The procedure used to quantify the JC-effect is the following:

- For all events, identify the channels with (ADC$-$pedestal) value >200, and the Beetle chips to which these channels belong.
- For these events, consider all channels of this Beetle chip with (ADC$-$pedestal) <40 and calculate the average (ADC$-$pedestal) value for these channels.
- This value corresponds to the magnitude of the JC-effect for a specific Beetle chip and a specific event.

In Fig. 4.14 the ADC values of the channels of one Beetle chip with (ADC$-$pedestal) <40 is shown, where at least one channel has (ADC$-$pedestal)>200. The negative tail indicates the shift to lower ADC values.

As explained previously, the purpose of calculating the magnitude of the JC-effect was to see if, once the shift in the ADC distribution due to this effect was corrected for, a visible channel-to-channel correlation would still remain. So the magnitude of the JC-effect was subtracted from the ADC value of all channels of that Beetle chip.

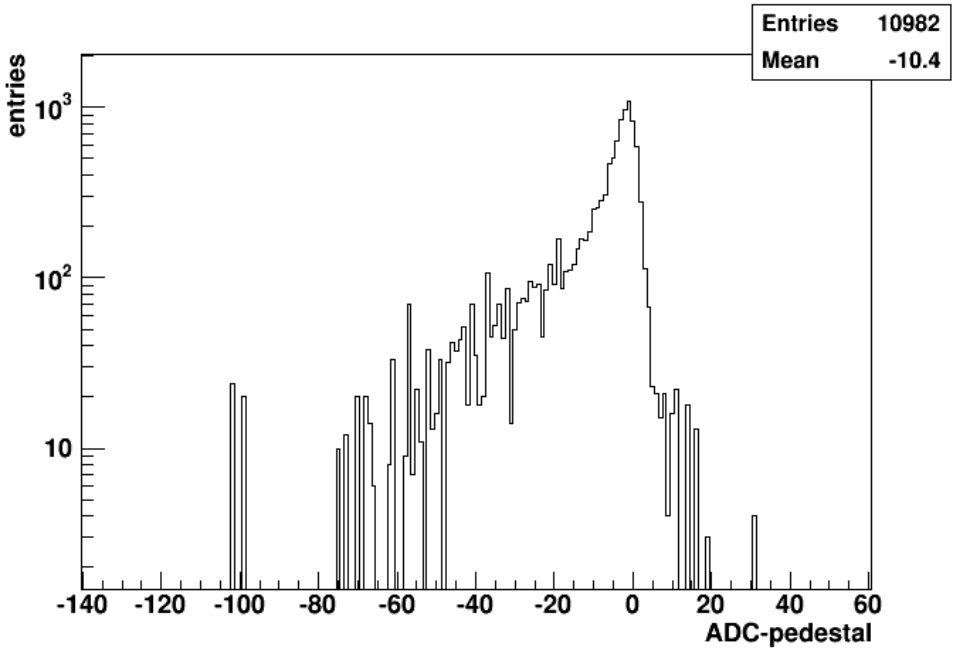

Figure 4.14: *Average ADC values of the channels for which (ADC−pedestal) < 40. Only the channels of Beetle chips for which at least one channel has (ADC−pedestal) > 200 are selected.*

In Fig. 4.15 the original (ADC−pedestal) spectrum of the analogue channels is depicted in blue, while the red curve indicates this spectrum after the JC-effect has been corrected for.

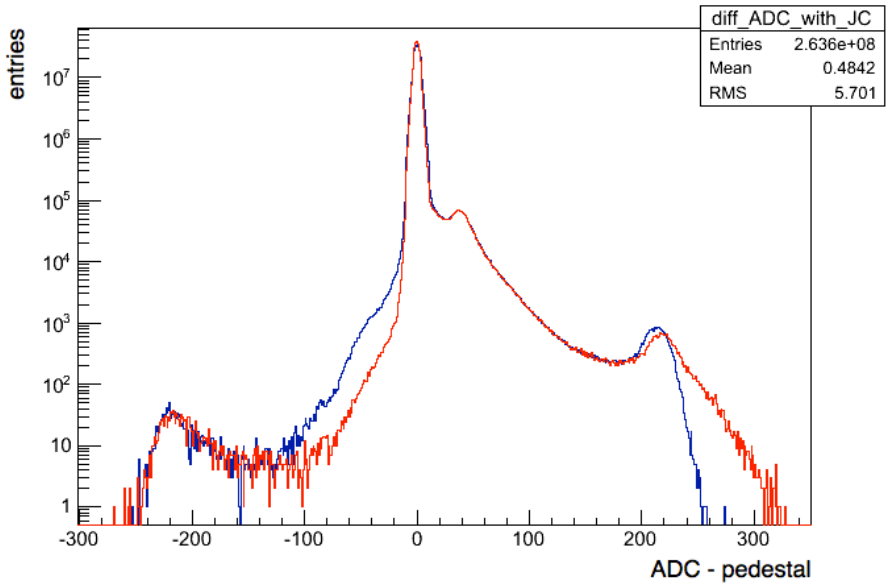

Figure 4.15: *(ADC−pedestal) distribution of all analogue channels for all sensors and all links before correcting for JC-effect (blue) and after correcting for JC-effect (red).*

Therefore, the difference between the two curves is due to the pedestal correction in Fig. 4.14. After correction there are less entries with a value in the range $[-100, -20]$. Similarly, there are more entries with a value larger than 260 ADC counts. The reason for which these two numbers of entries are not equal can be explained as follows. If, for example, there are two channels with a high ADC value, the ADC value of the other 126 channels of the same Beetle chip will be shifted towards lower values. When we correct for the JC-effect, the calculated magnitude of this effect is added to the (ADC−pedestal) value of all of the 128 channels of that Beetle chip. Therefore, there will be 126 less channels in the range $[-100, -20]$ and only 2 more channels with an (ADC−pedestal) value greater than 260.

It can also be seen in Fig. 4.16, which contains a zoom of the central region of Fig 4.15, that the peak of the distribution after the JC-effect has been corrected for is higher and narrower with respect to the peak of the uncorrected one, showing that after this correction there is a larger number of channels with an (ADC−pedestal) value around zero.

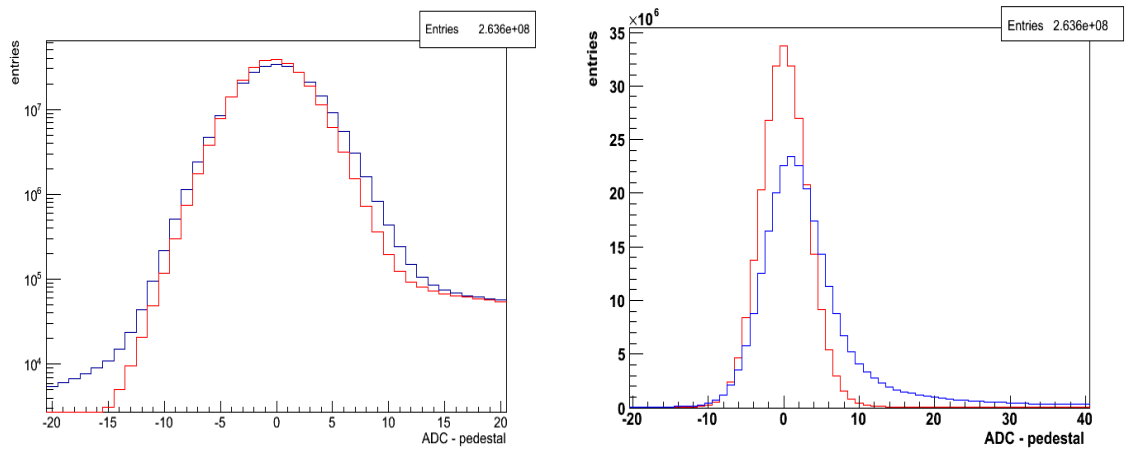

Figure 4.16: *Zoom of the peak of the logarithmic (left) and linear (right) ADC distribution of all analogue channels for all sensors and all links before correcting for JC-effect (blue) and after correcting for JC-effect (red).*

In order to see how much the shift in the ADC distribution is due to the JC-effect, the correlation between the (ADC−pedestal) value of strip-channel 0 and the ADC value of header bit 3 was examined after correcting for the JC-effect. In Fig. 4.17 the correlation between the ADC values of strip-channel 0 and header bit 3 is depicted before (left) and after (right) correction of the JC-effect. It can be seen that, after having corrected for the JC-effect, there are less entries in the range of (ADC−pedestal) values $[-100, -20]$. This region is circled in red in the figure for both low and high headers.

It is also interesting to observe that the slope of the distributions for low and high headers, before and after subtracting the JC-effect, remains unchanged. In fact, when calculated, both slopes are equal to about -5%. This indicates that the correlation between the (ADC−pedestal) value of strip-channel 0 and the ADC value of header bit 3 is independent of the JC-effect, as is the case for strip-channel 31 and header bit 0.

4.4 Summary of Results

In this section three different cases were chosen to study the behaviour of headers (and strip-channels) close to strip-channels with high signals. First, the correlation between high signals on

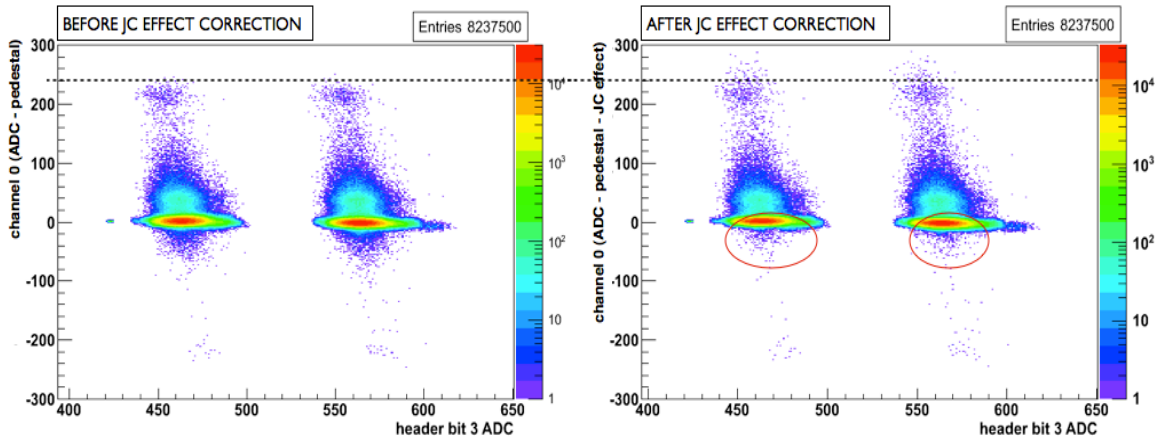

Figure 4.17: *Correlation between (ADC−pedestal) values of strip-channel 0 and ADC values of header bit 3 before (left) and after (right) correcting for the JC-effect.*

strip-channel 0 and the ADC value of header bit 3 was studied, both before and after correcting for the JC-effect, in order to determine if there is a dependence from this effect. Second, the correlation between high signals on strip-channel 31 and the ADC value of header bit 0 of the subsequent event was analysed. Having found that there is no dependence on the JC-effect, the second case was studied only before correction. Third, the correlation between a high signal on one strip-channel and the signal on the neighbouring strip-channel was also investigated for the various strip types of the ϕ-sensors.

Both for headers and strip-channels the effect of a high analogue value was examined for previous and next strip-channels or headers. The ratio between the (ADC−pedestal) value of header bit 3 and that of strip-channel 0 is approximately -5%, while in the case of header bit 0 and strip-channel 31 this ratio is about -4%.

For neighbouring (non-adjacent) strips on ϕ-sensors a typical correlation of -6% to -4% was observed between the (ADC−pedestal) value of one channel and the negative shift of the pedestals of neighboring channels. This study shows that the observed correlation results from cross-talk in the electronics and not from an effect on the detector. The magnitudes of the ratios for the correlation between headers and strips are comparable within the precision of the method of $1\% - 2\%$. In all cases, a high signal on one strip-channel causes the value of the 'neighbouring' header to be shifted towards lower values.

The ratio between the centers of the high and low header distributions was also measured for both header bit 3 and header bit 0. Again, a value of approximately -4% was measured for the distributions of header bit 3, while for header bit 0 this ratio is equal to about -6%. This is comparable to that calculated separately for all distributions. Also in this case, to lower header values are associated higher (ADC−pedestal) values of the 'neighbouring' analogue strip-channels, confirming that there is a correlation between the high signal on a strip-channel and the 'lower than average' value of the neighbouring header (or strip-channel).

The ratio between the (ADC−average) of header bit 3 and the (ADC− pedestal) value of channel 0 is also approximately -4% before and after correction of the JC-effect. This confirms that the correlation between these values is independent of the JC-effect, and must therefore have its origin in the readout chain.

The existing correlation between the (ADC−pedestal) value of consecutive headers and channels corresponds to about -5%. This is only observed between consecutive headers and channels and between non-adjacent strips, due to the fact that charge sharing between adjacent strips 'compensates' for this effect. For what concerns header channel cross-talk, a shift of -5%

in the (ADC−pedestal) value of an analogue channel will not have observable consequences for the reliability of the data. Instead, in extreme cases the channel header cross-talk can cause the value of a header bit to fall in-between the high and low header thresholds, leading to the presence of a pseudo header error in that event. As a result, the event will be considered 'faulty' and will not be included in the data sample processed for physics analyses. However, since the number of events that can be affected by this phenomenon is small compared to the total number of events in a data sample, there will be no consequence on the physics measurements obtained using VELO data.

Chapter 5

Study of Detector Effects via Negative Hits

While performing the analyses on run 81811 (70337) to study the correlation between the ADC values of header bit 3 (0) and channel 0 (31), a relative large occurrence of hits with (ADC−pedestal) < -150 ADC counts was observed. It was decided to study these "negative" hits separately, to try to understand what their source is. In Sec. 5.1 the distribution of these hits in different time samples for R and ϕ-sensors is discussed. A more detailed study is carried out in Sec. 5.2 on ϕ-sensors only, in order to determine if there is a correlation between the presence of a negative signal and its location on the sensor. In this section, three effects that are observed in correlation with the presence of a negative hit are discussed: a Beetle undershoot (Sec. 5.2.1), induction in the sensor due to drift of charge (Sec. 5.2.2), and negative signals generated by high signals (Sec. 5.2.3). In Sec. 5.3, the ratio of the number of negative hits observed in the inner and in the outer region of the sensor is compared with the ratio of the expected number of particle interactions. The results of these investigations are summarized in Sec. 5.4.

5.1 Negative hits in different time samples for R and ϕ-sensors

A dedicated study of negative hits was performed in run 70337, containing 19881 events. This dedicated run was obtained reading out 5 consecutive 25 ns time samples. In this run 5248 events were found with (ADC−pedestal)< -150 in at least one time sample of at least one channel. In Fig. 5.1 the distribution of these negative hits per analogue channel per time sample is shown. For this run, the collision between the LHC beams occurs in time sample 1. It can be seen that the negative hits are evenly distributed across the 32 analogue channels. Moreover, there are no negative hits observed in time sample 0.

In Fig. 5.2 the distribution of negative hits per link versus sensor number is shown. Of the total number of 5248 negative hits in all time samples together, 4282 hits occur on the ϕ-sensors (sensor number > 63) and only 966 hits on the R-sensors (sensor number < 42); in both cases evenly distributed over all links. Clearly, there are much less negative hits in R-sensors than in ϕ-sensors.
The distribution of negative hits in ϕ-sensors is further investigated in Sec. 5.2.

5.2 Negative hits in different sectors of the ϕ-sensors

The difference in the number of negative signals on R and ϕ sensors can be due to the different layout of the two types of sensors. Contrary to R-sensor strips, the strips on the ϕ-sensors

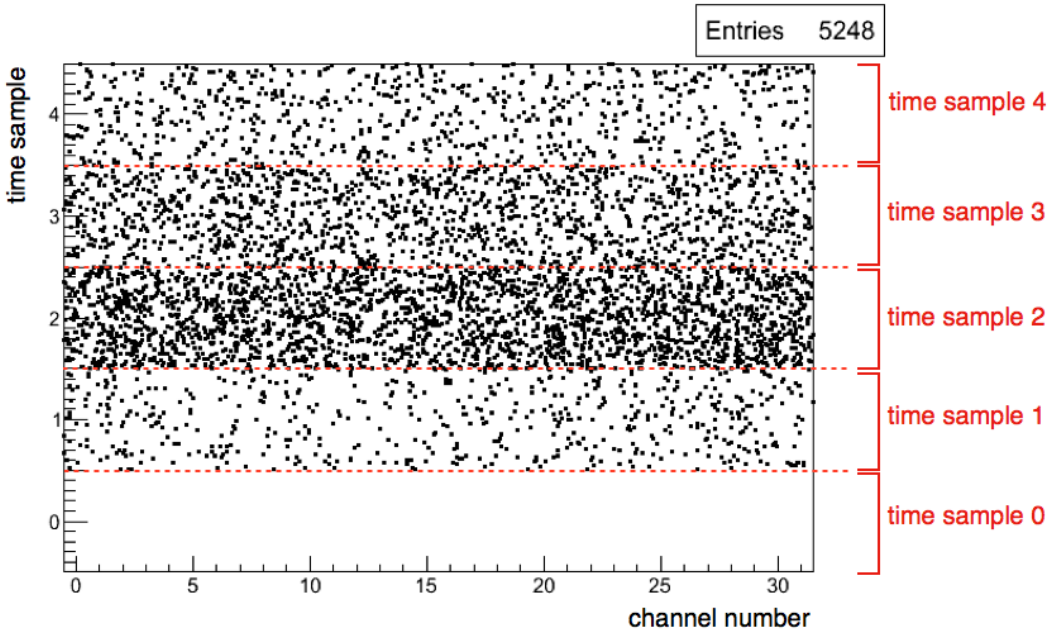

Figure 5.1: *Occurrence of negative ((ADC−pedestal)*< -150*) hits per time sample versus channel number. The dashed lines indicate the separation between the different time samples. No channel number structure is observed.*

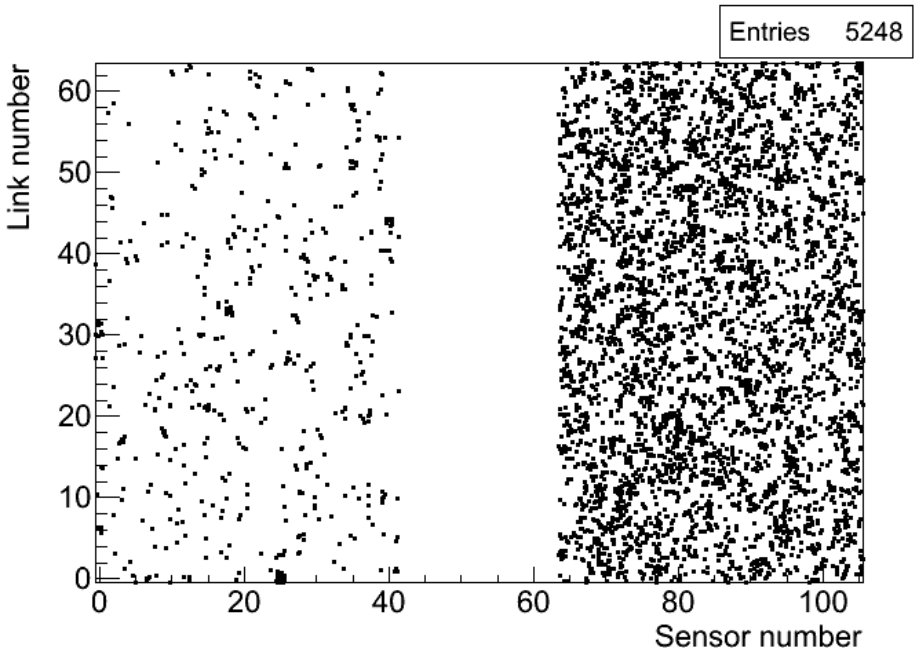

Figure 5.2: *Distribution of negative hits per link versus sensor number, where R-sensors* $\in$*[0,41], ϕ-sensors* $\in$*[64,105]. All time samples are added up.*

can be classified in two distinct categories: 683 inner and 1365 outer strips (see Fig. 3.2). In Fig. 5.3 a schematic cross section of a VELO ϕ-sensor is shown. The signal on the inner strips

is transferred to the read-out electronics through routing lines, which pass over outer strips to reach the edge of the sensor. The two traces are separated by a 3.5 μm thick SiO_2 layer and are each connected to an amplifier, which processes the input signal before it is sent to the shaper of the Beetle. Note that, since there are twice as many outer strips than inner strips, only half of the outer strips have an inner strip routing line running over them.

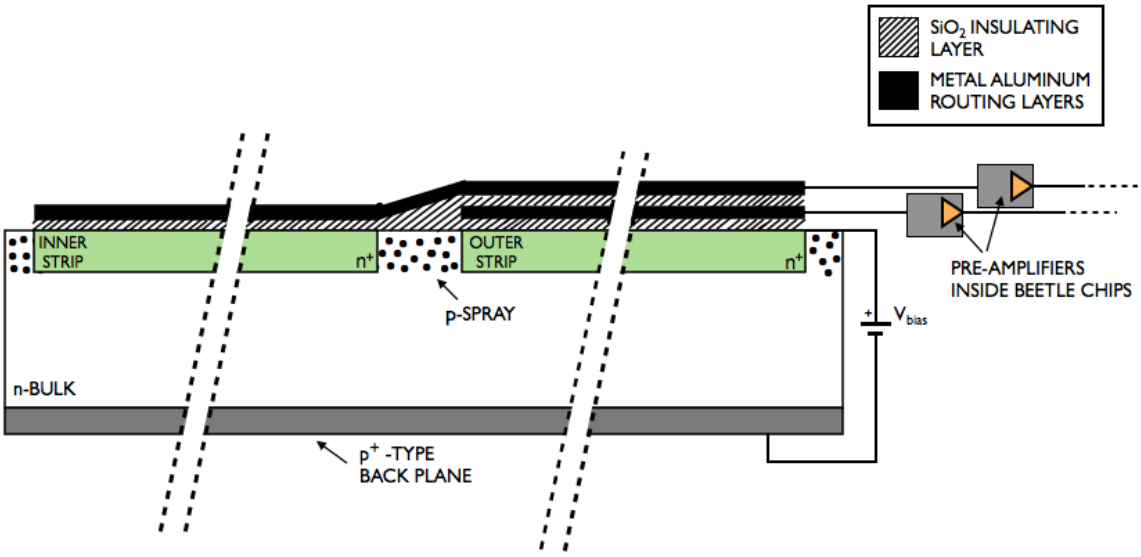

Figure 5.3: *Schematic cross section of a ϕ-sensor. The inner strips (left) are connected to the Beetle chip pre-amplifier via routing lines overlaid to outer strips (right). Only the pre-amplifiers of the Beetle chip are shown.*

The strips in the outer region of the ϕ-sensor can therefore be separated in two groups: outer strips with routing line and outer strips without routing line. By comparing the number of negative hits on the different strip types, it is possible to determine whether these hits have their origin in the sensor itself or that they are generated by the routing patterns.

For this study, events with a value of (ADC−pedestal)< -150 on at least one channel were selected. The channel numbering was reordered in order to match the strip numbering on the sensor. In this way, consecutive channels represent adjacent strips on the sensor. Out of a total of 5282 negative hits, 4282 were found on ϕ-sensors. Of these, 1848 hits (43%) were on inner strips, 2153 hits (50%) on outer strips with an overlaid routing line, and 281 hits (7%) on outer strips without an overlaid routing line.

These numbers illustrate that there is a correlation between the presence of a routing line and the occurrence of a negative hit on a strip. The nature of this correlation is further investigated below. Due to the number of instrumental effects that can affect the signals, it is not straightforward to identify the origin of the negative hits. Some events are very 'busy' and it becomes difficult to classify each effect independently. After visual inspection three effects were identified as the possible origin of negative signals: the Beetle undershoot, induction in the strips due to drift of charge in the sensor, and negative signals generated by high signals[1]. Each is discussed below and illustrated with an event plot, where inner strips are drawn in black, outer strips with an overlaid routing line in red, and outer strips without routing line in green. Negative signals due to the Beetle undershoot or induction in the sensor are rarely greater[2] than

[1]Signals greater than 150 ADC counts.

[2]In this chapter, the terms 'greater than' and 'less than' refer to the magnitude of a negative signal. For example, a signal of −200 ADC counts is considered greater than −150 ADC counts.

−100 ADC counts, while the latter effect generates negative signals with a magnitude larger than −150 ADC counts. However, the first two effects will be described for completeness.

5.2.1 The Beetle undershoot

When a sensor is hit by a particle, the charge deposited in the sensor is read out by the Beetle chip, which then requires some time to 'restore'. This results in a pulse shape that passes through negative (ADC−pedestal) values before stabilizing again around zero (see Fig. 5.4). It

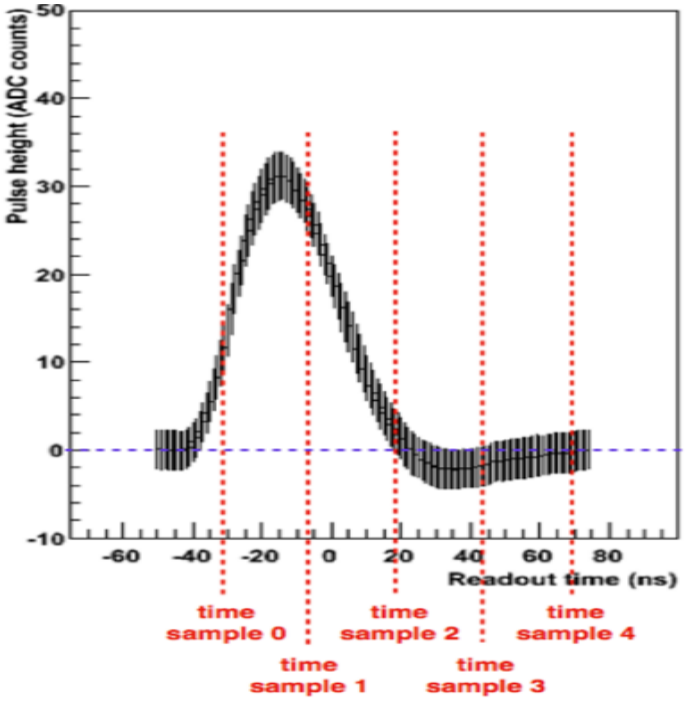

Figure 5.4: *Profile of the Beetle pulse shape. The dashed lines indicate the times of the subsequent sampling moments.*

can be seen that the maximum of the pulse shape corresponds to time sample 1, while in time sample 3 the Beetle undershoot is observed.

For each channel, the correlation between the negative hits in time samples 2 and 3 and the ADC values in other time samples was investigated, in order to see if the negative signals correspond to the Beetle undershoot. These two time samples were selected because of the large number of negative hits observed in them. In Fig. 5.5 the distribution of the (ADC−pedestal) values in all time samples, which correspond to negative hits ((ADC−pedestal)< -150) in time sample 2 and 3 are shown on the left and right, respectively.

The data with a negative hit in time sample 3 match the profile of the Beetle pulse shape, which suggests that the negative hits observed in this time sample could be due to a large undershoot of the Beetle. The concentration of high hits observed in time sample 1 correspond to the maximum of this pulse shape, which is consistent with the profile of a large observed signal.

The amplitude of the undershoot is typically about 1/8 of the maximum amplitude of the pulse (see Fig. 5.4). Many observed negative hits have an amplitude of about −200 ADC counts, which means that for these events we should observe hits of about 1600 ADC counts in time sample 1. The fact that the maximum observed values are around 220 ADC counts is due to the saturation of the Beetle chip. As was mentioned earlier, approximately 40 ADC counts

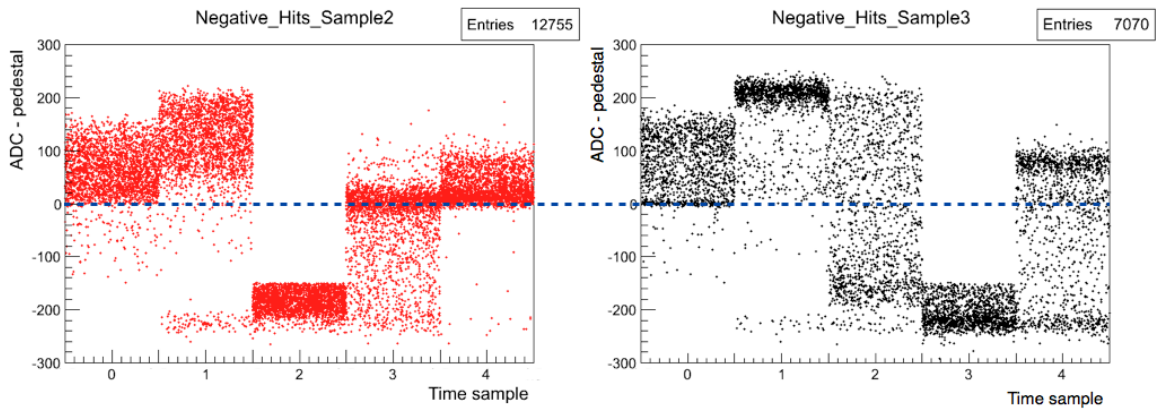

Figure 5.5: *Distribution of the (ADC−pedestal) values in all time samples, which correspond to negative hits ((ADC−pedestal)*< -150*) in time sample 2 (left) and 3 (right).*

correspond to 1 MIP, which is equal to approximately 24000 electrons. Therefore, values of 1600 ADC counts ($\simeq$ 1 million electrons) are much too high for the shaper circuit of the Beetle front-end to handle. A signal high enough to create an undershoot of −150 ADC counts would take long to disappear, meaning that such an undershoot should also be observed in time sample 4. An example of a regular Beetle undershoot event is shown in Fig 5.6. Here, the signal on an

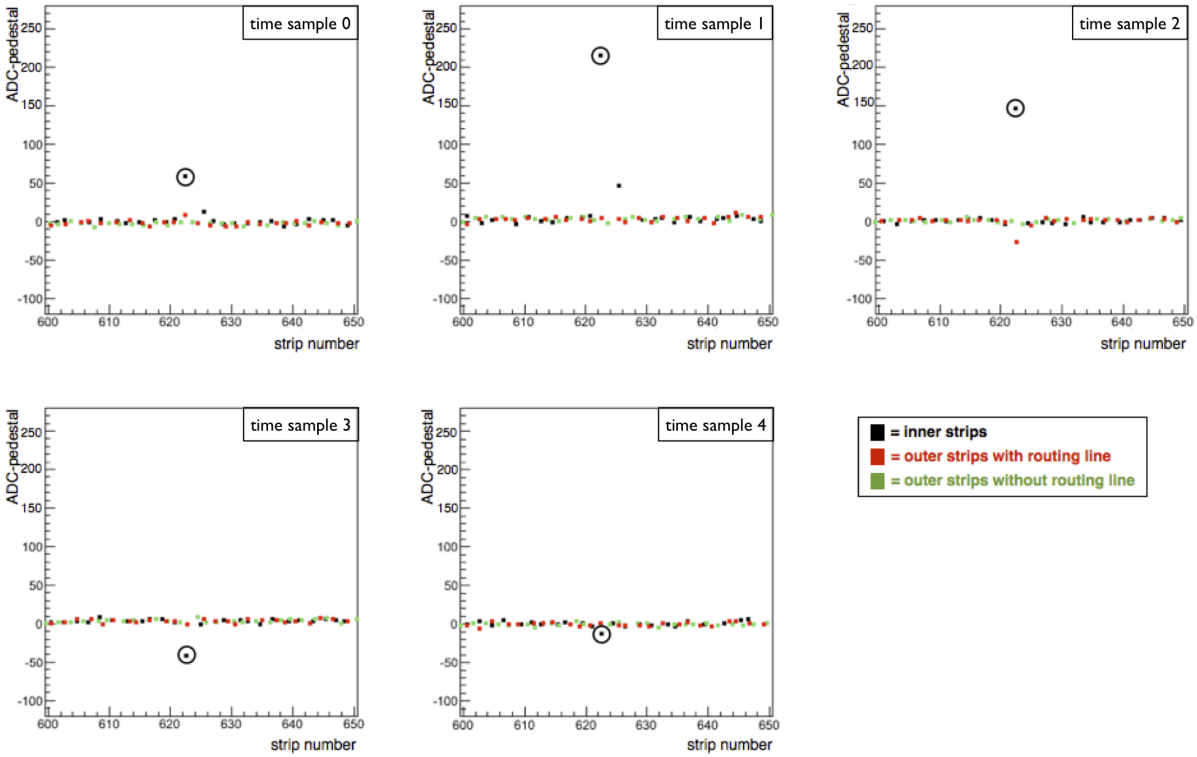

Figure 5.6: *Event in which a negative signal reflecting the Beetle undershoot is observed in time sample 3 on an inner strip. The evolution of the signal on this strip is shown for the 5 time samples, circled in black.*

inner strip (circled in black in the figure) begins to increase in time sample 0, reaches saturation in time sample 1, begins to decrease in time sample 2, is negative in time sample 3 and, finally, returns to zero in time sample 4.

The data with negative values in time sample 2 (left in Fig. 5.5) instead does not fit this pulse shape profile. In this case the ADC value drops abruptly between time samples 1 and 2 to larger than −150, indicating that the observed negative hits are caused by an effect different from the Beetle undershoot.

Although negative signals due to the Beetle undershoot were observed frequently in time samples 3 and 4 and do not occur in time samples 1 and 2, it is rare for the magnitude of these signals to be greater than −150 ADC counts. Consequently, the large number of hits with (ADC−pedestal)< -150 cannot be explained by this effect.

5.2.2 Induction in the sensor due to drift of charge

The second class of negative events consists of events in which the negative signals are likely caused by induction in the sensor [55]. An example of such an event is shown in Fig. 5.7. In the analysis, this type of negative signals was observed in approximately 30% of the cases, but the amplitude of these signals is never larger than −100 ADC counts.

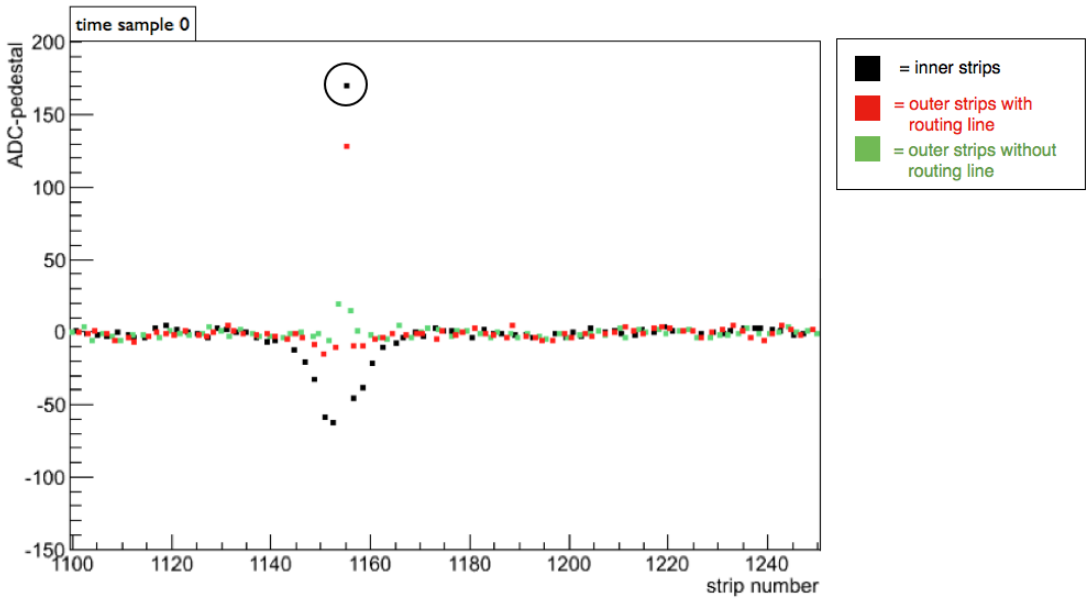

Figure 5.7: *Event in which a high signal present on an inner strip (encircled in black) simultaneously induces negative signals on the neighbouring inner strips. Only time sample 0 is shown.*

To investigate this effect in more detail the signal development in the sensor and the subsequent response of the electronics was evaluated. After a passing particle liberates charges in the bulk, these charges will start drifting and hence induce currents on nearby strips. For a moving point charge q, the induced current on a particular strip can be calculated by using the Shockley-Ramo theorem [56], [57], as:

$$i = q\vec{v} \cdot \vec{E}_0(\vec{x}) \tag{5.1}$$

where $\vec{v}$ is the instantaneous velocity of the charge q and $\vec{E}_0(\vec{x})$ is known as the *weighting field* of that strip, which describes the coupling between the strip and the charge. The shape of the induced pulse will depend both on the amount of charge and the position at which it is liberated, and can be bipolar. The measured signal will be given by the convolution of the induced current pulse and the response of the Beetle front-end chip (a CR-RC shaper), which has a peaking time of about 20 ns, and is sampled at 25 ns intervals.

The development of signals in a silicon sensor has been simulated with the Synopsys TCAD software package. For minimum ionizing particles the observed pulse is positive[3]. Strips that are further away from the liberated charge will see a bipolar signal which integrates to zero within the peaking time of the Beetle front-end.

However, if the liberated charge is very large (e.g. from slow and highly charged fragments from nuclear interactions) the drift field will be heavily distorted, which affects the drift time of the released charges. In that case the charge collection is slow and the observed signal can either be positive or negative at peak time, depending on where in the sensor the charge is liberated. It was observed in the simulation that if the charge is liberated at a distance of $\leq$50 μm from a strip, all neighboring strips see a signal after 25 ns which is opposite in polarity to that of the nearest strip. The amplitudes of these signals decrease with increasing distance to the point of interaction. The opposite polarity signals are mainly due to the fact that the weighting field of the neighboring strips has a direction opposite to that of the strip closest to the liberated charge.

In conclusion, the TCAD simulations confirm for specific conditions what has been observed experimentally, i.e. signals of opposite polarity are seen in adjacent strips in the same time sample. The amplitudes decrease with increasing distance to the hit strip.

5.2.3 Negative signals generated by high signals

Approximately 75% of the negative hits ((ADC−pedestal)<−150) are due to a different cause. They exhibit a pattern related to strips with overlaid routing lines as shown in Fig. 5.3: when there is a high signal on one strip, the (ADC−pedestal) value on one (or more) neighbouring strips of opposite type (i.e. outer strip with routing line if the high signal is on the corresponding inner strip, or inner strip if the high ADC is on an outer strip with routing line) is first 'pulled up' to a value comparable to that of the high ADC value and then drops below −150 ADC counts in the next time sample. Two time samples later this signal has returned approximately to the pedestal value. It therefore looks as if the negative (ADC−pedestal) value 'compensates' for the high signal in the previous time sample (i.e. the net detected charge is 0).

Events of this type were further classified, depending on the strip types having the high and the negative signal. Two sub-categories of events were found:

- Events for which there is a high signal present on one (or more) *inner strip(s)* which 'pulls up' the signal on one (or more) *outer strip(s) with routing line*. The signal on the outer strip becomes negative in the successive time sample.

- Events for which there is a high signal present on one (or more) *outer strip(s) with a routing line* which 'pulls up' the signal on the *inner strip* of which the routing line passes over the outer strip under consideration. The signal on the inner strip becomes negative in the following time sample.

In Figs. 5.8 and 5.9 examples of these two subcategories are depicted. For clarity, only time samples 1, 2 and 3 are shown. In Fig. 5.8, due to the high signal present on two inner strips, the signal on the corresponding outer strips with routing line is first pulled up (time sample 1),

[3]The polarity that is referred to here is the one as measured by the ADC and not that of the signal on the strip of the electron-collecting VELO sensors.

then drops to a negative value (time sample 2), and finally returns approximately to the baseline (time sample 3).

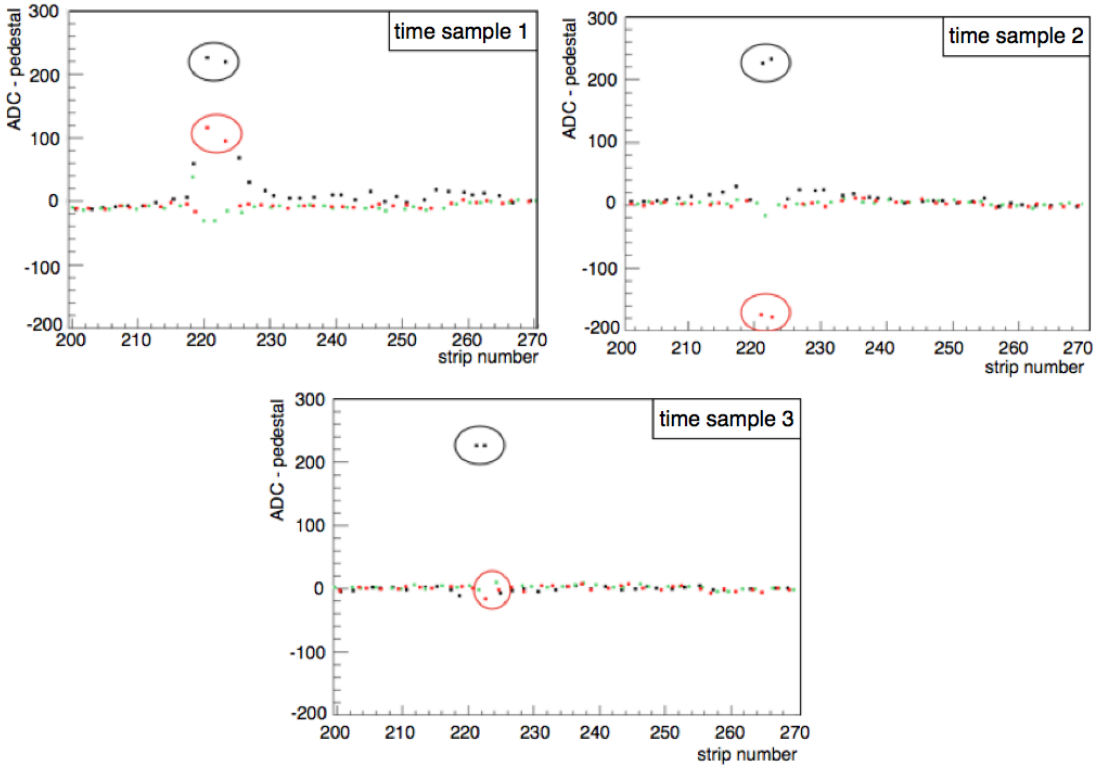

Figure 5.8: *In this event a high signal is present on two inner strips (black). For both strips, the presence of a high signal influences the corresponding outer strip with an overlaid routing line (red). The signal on each outer strip is 'pulled up' in time sample 1, becomes negative in time sample 2, and returns back at the pedestal value in time sample 3.*

In Fig. 5.9 the complementary situation is the case. This time, the high signal present on an outer strip with routing line causes the signal on the corresponding inner strip to be pulled up (time sample 1). The (ADC−pedestal) value of this signal is negative in time sample 2 and returns to the baseline in time sample 3.

Various events were found in which two or more negative signals, present on non-adjacent inner strips, are always generated by high signals present on outer strips with a routing line. In addition, no event was found in which there are combinations of negative signals on inner strips with only high signals on outer strips without a routing line.

A negative signal is present on an inner strip whenever there is a high signal on the outer strip that is coupled to the routing line of that inner strip, or viceversa. A possible explanation for this could be the following. If the deposited energy is very large, the Beetle chip pre-amplifier A_1 (see Fig. 3.5) saturates. This means that the excess charge will not enter A_1, but will instead accumulate at its input. Since the net charge at the input of A_1 must ultimately be zero, the charge in excess will be transferred, through capacitive coupling, to the routing line of the corresponding adjacent strip. The amplifier connected to the corresponding strip, A_2, will therefore see a same polarity signal, although no charge was actually deposited on this strip. The net charge 'seen' by A_2 must also be zero, so the magnitude of the negative signal must

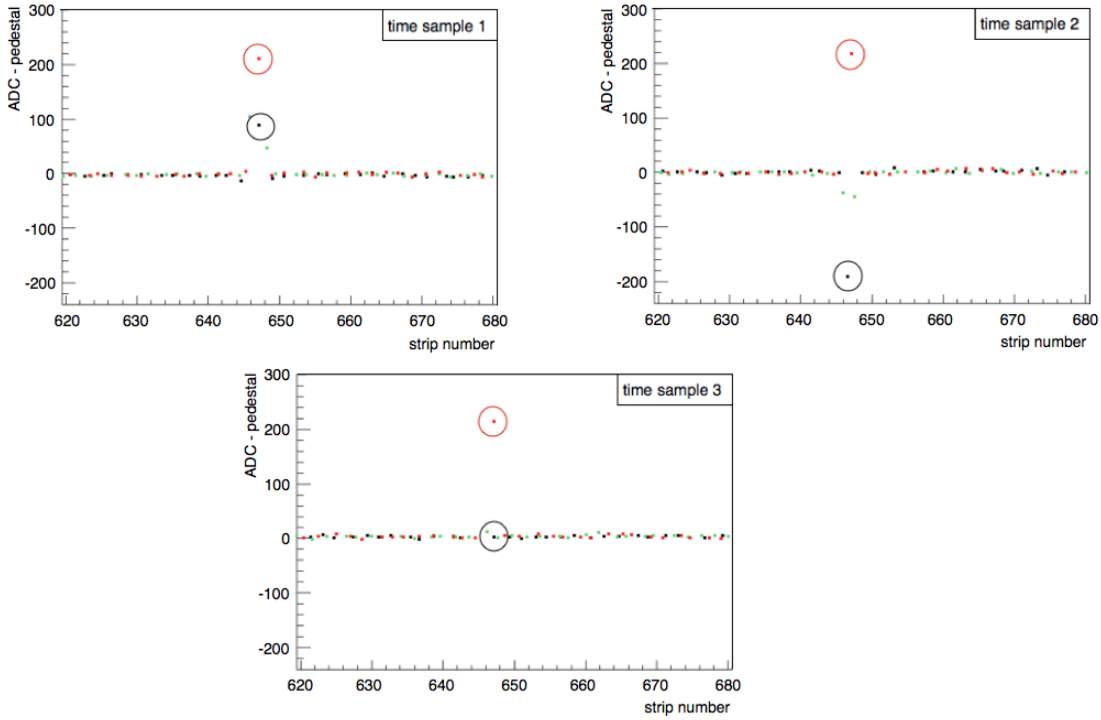

Figure 5.9: *Event for which the high signal on an outer strip with routing line (red) influences the corresponding inner strip (black). In time sample 1 the signal on the inner strip is 'pulled up', is negative in time sample 2, and returns to the baseline in time sample 3.*

be equal to the magnitude of the positive signal originating from capacitive coupling[4]. If the pre-amplifiers are too slow in draining the excess charge, a similar behaviour will be observed.

The correlation between the high signal on one strip and the observation of a negative signal on the neighbouring strip is still not completely understood. For instance, it is not clear why the negative signal is observed only in one time sample instead of remaining until the high positive signal on the hit strip disappears.

It is clear that the negative signals are correlated with the presence of a high signal on a corresponding strip. Moreover, the strip with the high signal and that with the negative one are always coupled via routing lines. It is not possible to conclude whether the observed negative signals are the result of coupling or if there is a *threshold effect* at a specific ADC value. This can possibly be resolved by performing a detailed simulation of the sensor with a Synopsys Technology Computer Aided Design (TCAD) software in combination with a Simulation Program with Integrated Circuit Emphasis (SPICE) simulation of the Beetle front-end. However, these studies are beyond the scope of this thesis.

[4]It must be noted that signals are only sampled at 25 ns intervals, meaning that the sampled peak amplitude could be different from the actual peak amplitude.

5.3 Occurrence of Negative Hits in the Inner Region of a ϕ-sensor

As was described in the previous section, 75% of the observed negative hits are associated with the presence of a signal greater than 150 ADC counts on the corresponding strip of opposite type[5]. These high signals are most likely due to nuclear interactions in the 300 μm thick silicon of the VELO sensors. On average, about 3000 clusters are observed in the VELO per event. Each cluster is generated by a particle crossing the 300 μm thick silicon. This means that the total material transversed in an event is equal to 3000×300 μm = 90 cm. Since the nuclear collision length in silicon is equal to 30.2 cm, on average three nuclear interactions are expected in the VELO for each event. By comparing the observed number of negative hits in one region of the sensor with the expected number of particle interactions in the opposite region, it is possible to verify the correlation of the occurrence of the negative hits with nuclear interactions in the sensor.

The expected number of particle interactions in the inner and outer region of a VELO ϕ-sensor is found to first approximation by integrating the expected particle flux over the surface area of each region. This particle flux is dominated by particles produced in primary interactions and depends on the radial position, r:

$$\text{Particle Flux} \propto \text{N} \times \text{r}^{\alpha} \tag{5.2}$$

where $\alpha \in [-2, -1.8]$ [34], and typical values of N are:
N = 0.6 particles/cm^2 for a minimum bias event,
N = 2.0 particles/cm^2 for a b-event.

In Fig. 5.10 the particle flux distributions for neutrons and charged particles are plotted. It can be seen that the contributions from neutrons are negligible with respect to the flux of charged particles.

The active inner region of a ϕ sensor extends from 0.8 cm to 1.72 cm, while the outer region goes from 1.73 cm to 4.2 cm. The total number of particle hits in each region is determined by integrating the particle flux over these two areas:

$$\int_{r_1}^{r_2} N \cdot 2\pi r \cdot r^{\alpha}\, \mathrm{d}r \tag{5.3}$$

To account for the uncertainty in the radial dependence of the particle flux, this calculation was done for $\alpha = -1.9 \pm 0.1$. For the inner region the expected number of particle hits amounts to (0.778±0.013)×N. For the outer region the expected number of hits amounts to $\left(0.980^{+0.103}_{-0.093}\right)$ ×N.

It has been argued that the presence of a negative hit on an inner strip is directly correlated with the presence of a very high signal on the corresponding outer strip (see Sec. 5.2.3). If this holds, the ratio between the expected number of particle interactions that leave a large energy deposition on one or more strips in the outer and inner region of the sensor should be comparable to the ratio between the observed number of negative clusters in the inner and outer region of the sensor. The first, R_{Flux} is given by:

$$R_{Flux} = \frac{\textit{Expected Particle Interactions in the Outer Region}}{\textit{Expected Particle Interactions in the Inner Region}} = 1.26^{+0.11}_{-0.10}$$

where the uncertainties are due to the error on the radial dependence. The latter, R_{Obs}, is given by:

$$R_{Obs} = \frac{\textit{Observed Negative Hits in the Inner Region}}{\textit{Observed Negative Hits in the Outer Region}}$$

[5] An inner strip if the negative signal is observed on the corresponding outer strip with overlaid routing line, and vice versa.

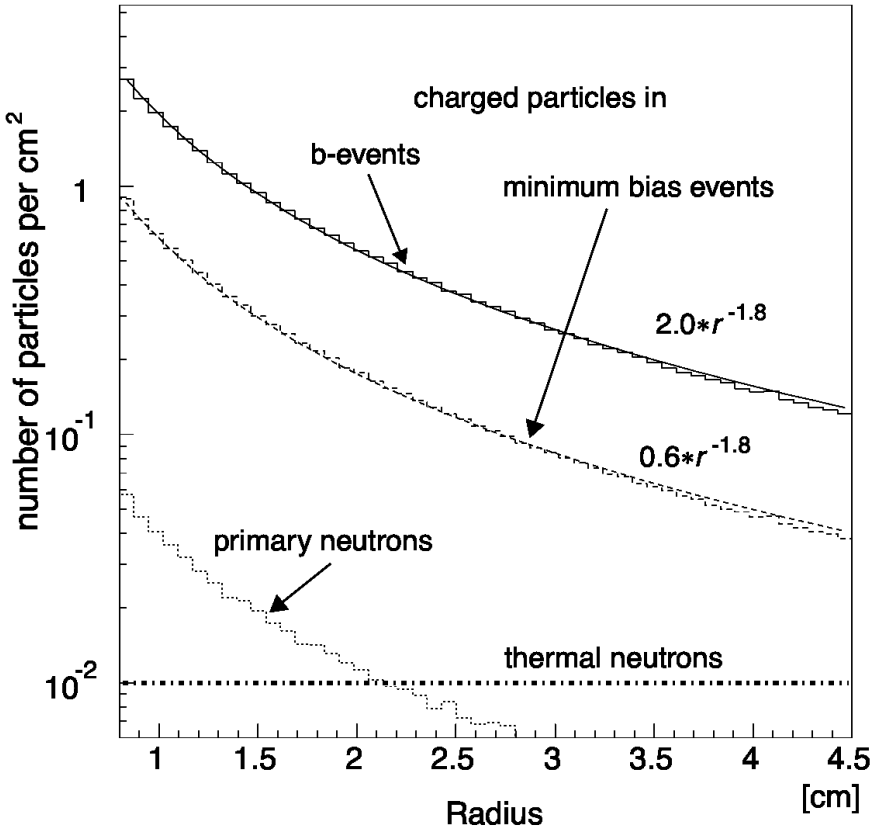

Figure 5.10: *Particle flux distributions for neutrons and charged particles. Here, the dependence on the radial distance from the center of a VELO sensor is taken as* $r^{-1.8}$.

In this analysis, a total of 1848 negative hits were observed in the inner region and 2434 in the outer region. The number of negative hits in the inner and outer regions for each time sample is listed in Table 5.1.

Time Sample	Inner Region	Outer Region
0	0	0
1	68	141
2	939	1247
3	556	679
4	284	367
total	1848	2434

Table 5.1: *Number of negative hits on the inner and outer region of a ϕ-sensor, per time sample.*

The largest number of negative hits was observed in time samples 2 and 3, so the data of these time samples was analysed in more detail. Since the results for the two time samples are comparable and time sample 2 has larger statistics, only the analysis of time sample 2 data will be described.

In time sample 2, 939 negative hits were observed in the inner region and 1247 in the outer region, giving:

$$R_{Obs} = 0.75 \pm 0.03 \tag{5.4}$$

where the uncertainty is statistical.

When calculating the number of candidate particle interactions in the outer region that can generate a negative hit in the inner region, it must be taken into account that negative signals on inner strips are only observed when the high signal is present on the corresponding outer strips with an overlaid routing line. Therefore, only half of the strips on the outer region will be able to 'induce' a negative signal. The number of candidate particle interactions in the outer region should then be divided by a factor $\lambda \in [1, 2]$, depending on the number of contributing strips in the outer region, yielding an effective R_{Flux}^{Eff}:

$$R_{Flux}^{Eff} = \frac{Expected\ Particle\ Interactions\ in\ the\ Outer\ Region/\lambda}{Expected\ Particle\ Interactions\ in\ the\ Inner\ Region}$$

In case only 1-strip clusters would be present, λ would have the value 2. In order to calculate this correction factor, further analysis was performed to identify and catalogue the clusters on the outer strips corresponding to inner strips with a negative signal. The frequency of negative hits was studied up to 9-strip clusters. For this purpose, only negative hits in the inner region were analysed.

In Table 5.2, the number of the different types of clusters in the outer region associated to negative hits in the inner region is listed[6].

Type of Cluster	Count	Frequency (%)
1-strip	145	16
2-strip	244	27
3-strip	168	19
4-strip	110	12
5-strip	84	9
6-strip	46	5
7-strip	24	3
8-strip	26	3
9-strip	18	2
$\geq$ 10-strip	39	4

Table 5.2: *Frequency of different types of clusters observed on outer strips associated to negative clusters on the corresponding inner strips, in time sample 2.*

For 1-strip clusters λ=2, while for other clusters the value is not clear. Although all of the 1-strip clusters were identified as outer strips with a routing line, these clusters represent only 16% of the total number of clusters that were observed. For this reason, the correction factor that needs to be applied to the expected number of candidate interactions on the outer region will then have a value between 1.16 and 2, implying that R_{Flux}^{Eff} has a value between $0.63^{+0.06}_{-0.05}$ (for λ=2) and 1.09±+0.09 (for λ=1.16), in agreement with $R_{Obs} = 0.75 \pm 0.03$.

For what concerns clusters made up of 2 or more strips, it was not possible to distinguish unambiguously the strip with the largest energy deposition. This is due to the saturation of the Beetle chip, which corresponds to an (ADC−pedestal) value of approximately 220 ADC counts.

As an extreme example, an event in which a negative hit was observed on the inner region of ϕ-sensor 93 and a high signal is present in the outer region corresponding to the inner strip with the negative signal. In the outer region of this sensor a highly energetic cluster formed by 26 strips is present. The signal on the outer strip that corresponds to the negative inner strip is equal to the ADC value at which the Beetle saturates, as are the ADC values of its neighbouring strips.

[6]It must be mentioned that, for approximately 4% of the events under consideration, a negative signal was observed on more than one adjacent inner strip.

In Fig. 5.11 the event display for this event is shown. In the figure, both the ϕ and the R-sensor (mounted back-to-back) are shown. The strips with a hit on the ϕ-sensor are depicted in blue, while for the R-sensor the strips are shown in red.

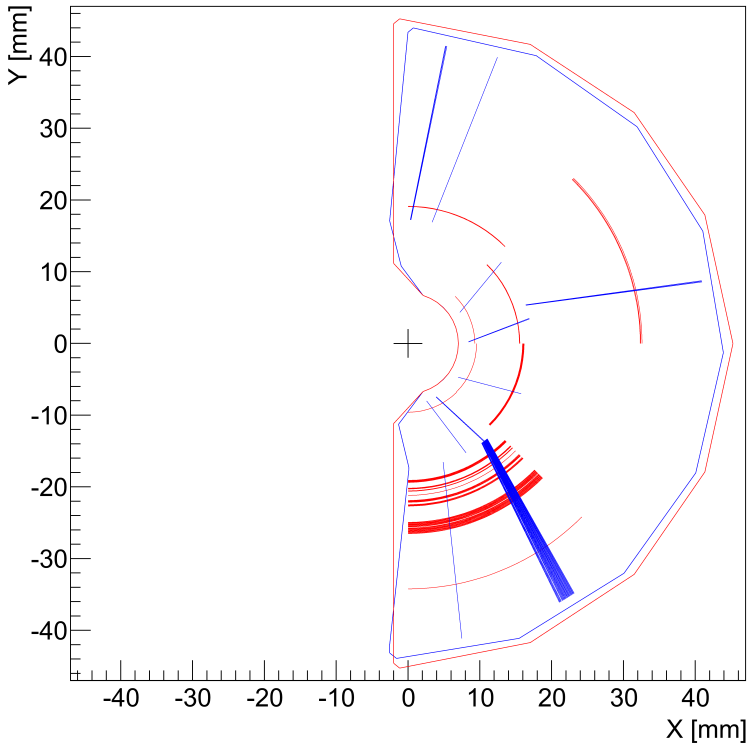

Figure 5.11: *Event display showing the strips with a hit on the ϕ (blue) and R-sensors (red) of the selected event. Each of the lines represent clusters with high signals formed by a number of strips that vary from one to four.*

In this particular module, the ϕ-sensor is mounted closer to the interaction region as the R-sensor, and it is interesting to see that a 23-strip cluster with a large signal is present also on the succeeding R-sensor, and in the same area of the sensor. This observation makes it plausible that a nuclear interaction takes place in the ϕ-sensor and, as a consequence of the large energy released in this interaction, there is a large energy deposition on the R-sensor that is mounted closely behind it.

5.4 Summary of Results

Negative signals were observed on both R and ϕ-sensors. Since approximately 80% of these signals were seen on ϕ-sensors, further analysis was performed on these sensors. Moreover, out of the total number of negative hits observed on ϕ-sensors, 43% were observed on inner strips, 50% on outer strips with an overlaid routing line and 7% on outer strips without routing line. Therefore, there seems to be a correlation between the presence of a routing line and a negative signal on a strip.

The origin of the negative signals is not easy to identify due to a number of instrumental effects that can play a role, such as the saturation of the Beetle chip. Nevertheless, three effects were identified as the possible sources of the negative signals: the Beetle undershoot, induction

in the sensor and negative signals generated by high signals present on strips connected via a routing line. The third effect that was identified is at the origin of approximately 75% of the observed negative hits. A correlation exists between signals of strips that are connected via a routing line such that, when there is a high signal on one strip, the (ADC−pedestal) value of one or more 'adjacent' strips of opposite type is first pulled up to a value comparable to the high signal and then drops to a negative value in the next time sample. It was observed that negative signals on inner strips are generated only if the high signal is present on the corresponding outer strips with an overlaid routing line. This is confirmed by the study of 1-strip clusters and from the fact that, if a negative signal is present on more than one inner strip, the associated cluster in the outer region is always observed in-between the outer strips that correspond to the first and last inner strip with a negative signal. The coupling between strips occurs via the routing line, and can well be due to the saturation of the amplifier connected to the strip with the signal.

To verify the existence of a correlation between negative hits in the inner (outer) region and high hits in the outer (inner) region, the expected number of particle interactions in the two regions of the sensor was calculated, leading to $R_{Flux} = 1.26^{+0.11}_{-0.10}$, where the errors account for the uncertainty on alpha: $\alpha = -1.9 \pm 0.1$. The expected number of interactions in the outer region must be divided by the correction factor $\lambda \in [1.16, 2]$ (only for 16% of the clusters it was possible to distinguish unambiguously that the negative hit in the inner region was generated by a high signal on an outer strip with an overlaid routing line), was then introduced. This leads to R^{Eff}_{Flux} having a value between 1.09 ± 0.09 and $0.63^{+0.06}_{-0.05}$, in agreement with the ratio between the observed number of negative hits in the inner and outer region, $\mathrm{R}_{Obs} = 0.75 \pm 0.03$ and indicating that the negative signals have their origin in the coupling of an outer strip with an inner strip via the routing line of the inner strip.

Chapter 6

$B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays

Hadronic decays of B_s mesons can be exclusively reconstructed without any missing momentum and are therefore well suited to study time-dependent decay rates. The hadronic decay of $B_s^0 \to D_s^- \pi^+$, with a branching ratio measured to be $(3.04 \pm 0.19(\text{stat.}) \pm 0.23(\text{syst.}) \, ^{+0.18}_{-0.16}(f_s/f_d)) \times 10^{-3}$ [58], is a self-tagging decay and can be used to measure the $B_s^0 - \bar{B}_s^0$ oscillation frequency. The topologically similar $B_s^0 \to D_s^\mp K^\pm$ decay, with a smaller branching ratio of $(1.97 \pm 0.18(\text{stat.}) \, ^{+0.19}_{-0.20}(\text{syst.}) \, ^{+0.11}_{-0.10}(f_s/f_d)) \times 10^{-4}$ [58], is not self-tagging and allows to measure the CP-violating angle γ, as was discussed in Chapter 1.

An unbiased determination of γ requires knowledge of the following experimental aspects: the decay time-dependent efficiency, referred to in general as time-acceptance, the decay time resolution of the reconstructed decays, the presence of backgrounds and the mistag probability of the flavour tagging procedure. These experimental aspects are discussed in this chapter. In Sec. 6.1 the data sample used and the event selection procedure are described. A generic description of the probability density function (pdf) used to model the proper time distribution and to extract the physics parameters is given in Sec. 6.2. The individual experimental ingredients of this pdf, the decay time resolution, detector acceptance and the background subtraction via the sFit method are described respectively in Sec. 6.3, 6.4 and 6.5. Lastly, the flavour tagging performance for $B_s \to D_s h$ decays is discussed in Sec. 6.6. Here, h represents the bachelor hadron, which is a kaon or a pion. In the remainder of the text we suppress the explicit notation of the final state particles, unless it is relevant, to include both charge conjugated decays.

The decays $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ are topologically identical. The topology of the decay $B_s^0 \to D_s^- h^+$ when $D_s^- \to K^+ K^- \pi^-$ is shown schematically in Fig. 6.1, where the final state h can be a pion or a kaon.
The B_s mesons are produced at the primary vertex from pp interactions, and on average decays after traveling a distance of about 1 cm. It decays at the secondary vertex, in the decay modes under study, giving origin to two final states that differ only by the bachelor particle, i.e. the particle which is produced in association with the D_s. After flying typically 6 mm the D_s meson can decay into two charged kaons and one charged pion. A $B_s \to D_s h$ decay is thus reconstructed by identifying a bachelor particle as a pion or a kaon, by determining the momenta of all final state particles to reconstruct the D_s and B_s invariant mass, and by requiring non-zero decay distances.

6.1 Event Selection

The event samples of $B_s \to D_s h$ candidates are obtained by a sequential selection procedure that includes the trigger, the stripping, and the offline selection. Since the $B_s \to D_s \pi$ is used as a tagging calibration channel for $B_s \to D_s K$, the selection procedure of the two decays is kept identical, apart from a final PID requirement on the bachelor particle. The event selection

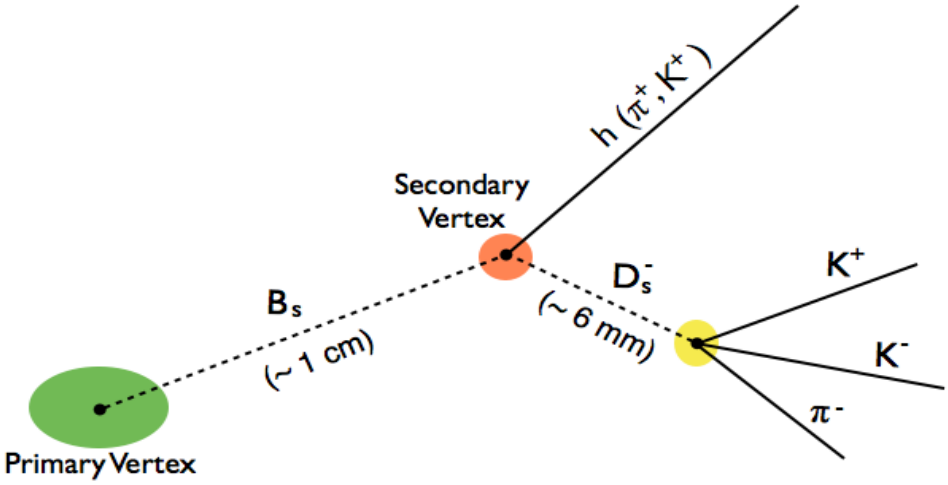

Figure 6.1: *Topology of a $B_s \to D_s h$ decay.*

that is applied in this analysis is based on that used for a previous analysis of $B_s \to D_s\pi$ and $B_s \to D_s K$ decays (see [29] for details).

The sample used in the simulation corresponds to $B_s^0 \to D_s^-\pi^+$ data corresponding to 2011 run conditions (MC11a), containing 12 million $B_s^0 \to D_s^-(K^+K^-\pi^-)\pi^+$ events, in which at least one $B_s^0 \to D_s^-(K^+K^-\pi^-)\pi^+$ decay is present in the detector acceptance. Calibration studies of the mistag probability are performed using a dedicated $B_s \to D_s\pi$ sample of data collected at LHCb in 2011, with approximately 1 fb^{-1} of proton-proton collisions at the LHC. These studies are described in Chapter 7.

Trigger

In the trigger, all events are selected by the topological trigger line `1TrackAllL0` in HLT1 and by the 2-, 3-, 4-body `TopoBBDT` lines in HLT2 [59]. First, HLT1 performs a partial event reconstruction using tracks that caused the L0 trigger to fire. Then, HLT2 looks for secondary vertices and performs a complete event reconstruction, leading to the selection of specific decay modes and event topologies. In order to be accepted an event is required to be of type `TOS` (Triggered on Signal), meaning that the same particles that fired the trigger should be used in the offline reconstruction to select the event.

Stripping

In the offline preselection, called stripping, kinematic pre-selection criteria are applied in order to reconstruct $B^0 \to D\pi$, $B_s^0 \to D_s\pi$, $B^0 \to DK$ and $B_s^0 \to D_s K$ decays. The $B_s \to D_s\pi$ candidates are selected by the `StrippingB02DPiD2HHHBeauty_ _2CharmLine` stripping line by assigning the pion mass hypothesis to the bachelor track, while the $B_s \to D_s K$ candidates are taken from the similar `StrippingB02DKD_ _2HHHBeauty2CharmLine` stripping line. In the stripping procedure events are pre-selected by applying loose cuts on the kinematics and distances from the primary interaction vertex of the final state particle. Also, selection criteria are applied to reconstructed D or D_s candidates. The list of kinematic and geometric cuts performed in the stripping is given in Table 6.1. In the table, the impact parameter of the track is indicated as IP, the distance of closest approach of the track with respect to a primary vertex (PV) is indicated as DOCA, and $IP\chi^2$ is defined as $IP^2/\sigma(IP)^2$.

Charged particles used to form b-hadrons
track χ^2/ndf < 4
p_T >100 MeV/c
p >1 GeV/c
IP $\chi^2 > 4$
D-candidates before vertexing
$\Sigma_i\ p_T^i > 1.8$ GeV/c , with i=daughters
maximal DOCA from PV < 0.5 mm
mass within 100 MeV/c^2 of nominal D^+ or D_s mass
D-candidates after vertexing
vertex χ^2/ndf <10
minimal vertex separation χ^2 from any PV > 36

Table 6.1: *List of selection criteria applied in the stripping of $B_{(s)} \to D_{(s)}h$ events. The term DOCA indicates the distance of closest approach of the track.*

The variables listed above, on which the stripping selection criteria are applied, are subsequently used as input variables to a boosted decision tree (BDT), to reduce the output rate of the data. A decision tree is a binary structured classifier, in which yes/no decisions are performed on a set of variables until a stop criterion is reached. The path down the tree to the leaf node represents an individual cut sequence that selects signal or background depending on the type of leaf node. A decision tree is trained by using a set of known training events, and the results are measured using a separate set of known testing events.

The boosting of a decision tree (BDT) consists in deriving different decision trees from the same training sample by reweighing events, and then combining them to form a classifier. For the boosting of a decision tree, if a training event is misclassified, i.e. a signal event is classified as background or vice versa, then the weight of this event is increased and a new tree is formed. Each decision tree is optimized such that the signal purity of the sample is maximized. The power of the cuts that are applied is represented by a loss function that has a stationary point at the point in which the cuts are optimal. For the preselection of $B_s \to D_s h$ events, the output of this BDT is required to be larger than 0.05, which is a loose cut that allows to keep basically 100% of the signal decays.

Offline Selection

After the stripping process is completed, an offline selection is applied to the events. Here, additional independent kinematic and geometric criteria are applied to the B_s and D_s candidates. In Table 6.2, the cuts applied during the offline selection are given. The first line indicates that the offline reconstructed particles should also be used in the trigger selection, hence the name Trigger On Signal (TOS).

A gradient boosted decision tree[1] (BDTG) training is then performed on the sample in a data-driven way. This means that both the training sample that represents signal as well as the training sample that represents combinatorial background are taken from the original data sample. The first consists of a weighted data sample, with events in the mass range $5310 < m(B_s^0)$ <5430 MeV/c^2, created using the *sPlot* technique (which will be described in Sec. 6.5). The latter is selected from the upper B_s^0 mass sideband, $m(B_s^0) > 5445$ MeV/c^2. The variables used as input to the BDTG are listed in Table 6.3.

[1]In gradient boosting a prediction model is defined as the combination of weak prediction models, representing different sets of cuts. The final prediction model is determined by moving along a differentiable *loss function* until the optimal ensemble of cuts is found.

Description	Requirement
trigger topo	TOS
D_s^+ mass	[1940,1990] MeV/c^2
D_s^+ lifetime (wrt. B_s^0)	> 0 ps
D^+ veto:	
PIDK of same charge K	> 10, or
$\lvert m(K^+K^-)$ - 1020 MeV/c$^2 \rvert$	< 10 MeV/c^2, or
D_s^+ under D^+ hypothesis	not in [1840,1900] MeV/c^2
Λ_c veto:	
p veto, same charge K	PIDK - PID$p > 5$, or
D_s^+ under Λ_c hypothesis	not in [2250,2320] MeV/c^2

Table 6.2: *Kinematic selection requirements. The cuts used in the D^+ and Λ_c vetoes are applied to $D_s^+ \to K^+K^-\pi^+$ candidates. PIDK (PIDp) represents the logarithm of the ratio between the particle's probability of being a kaon (proton) and the probability of it being a pion. TOS refers to Trigger on Signal.*

B_s candidate
DIRA OWNPV
IP χ^2
radial flight distance
vertex χ^2/ndf
lifetime vertex χ^2/ndf
bachelor
minimum IP χ^2
p_T
cos θ
D_s candidate
DIRA ORIVX
DIRA OWNPV
radial flight distance
vertex χ^2/ndf
D_s children
minimum IP χ^2
minimum p_T
bachelor and D_s children
maximum track ghost probability

Table 6.3: *BDTG input variables. The variable* $\cos(\theta) = \dfrac{\vec{x}\cdot\vec{p}}{|\vec{x}||\vec{p}|}$ *gives a measurement of the angle between the direction of the particle's momentum $\vec{p}$ and flight vector. The variable DIRA OWNPV= $\cos(\theta)$ with $\vec{x} = \vec{x}_{DV}$ - $\vec{x}_{PV}$, where DV stands for decay vertex and PV for primary vertex. For DIRA ORIVX instead we have $\vec{x} = \vec{x}_{DV}$ - $\vec{x}_{origin}$, where origin stands for the production vertex of the D_s meson.*

The final response is a continuous number that comes from the combination of the final responses of the individual trees. In order to be included into the data sample $B_s^0 \to D_s^-\pi^+$ events are requested to have a BDTG response that is larger than 0.5. In Fig. 6.2, the distribution of

the BDTG response variable is shown for the test sample. The individual signal and background distributions are depicted in red and blue, respectively.

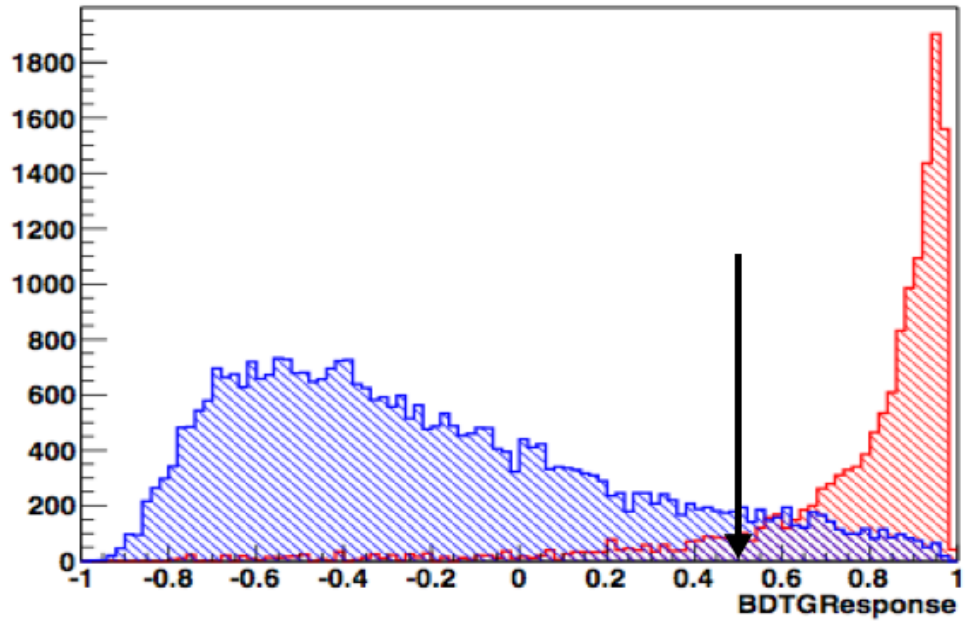

Figure 6.2: *Distribution of the BDTG response for the test sample. Signal events are shown in red, background events are shown in blue.*

Finally, $B_s \to D_s\pi$ and $B_s \to D_sK$ decays are disentangled by applying additional PID cuts on the bachelor track. The cut on the PID of the bachelor and that on the output of the BDTG are optimized with respect to the D_sK signal yield significance:

$$S_{D_sK} = \frac{N_{D_sK}}{\sqrt{N_{D_sK} + N_B}}.$$

To select $B_s \to D_s\pi$ decays the requirement PIDK < 0 must be met, while to select $B_s \to D_sK$ decays it is optimal to require PIDK > 10.

In Fig. 6.3 the results of a simultaneous fit to the mass distribution of $B_s^0 \to D_s^-\pi^+$ candidates and background contributions is shown [29]. The backgrounds that contribute to the $B_s^0 \to D_s^-\pi^+$ candidates can be divided in five groups, and are depicted in of Fig. 6.3 in different colours. The first group of background events consists of partially reconstructed final states, such as $B_s^0 \to D_s^-\rho^+$, $B_s^0 \to D_s^{*-}\rho^+$ and $B_s^0 \to D_s^{*-}\pi^+$, and are depicted in light blue in the figure. The second group consists of the fully reconstructed $B^0 \to D^-\pi^+$ events (yellow). The third group consists of partially reconstructed $\bar{\Lambda}_b^0 \to \bar{\Lambda}_c^-\pi^+$ events (red). The fourth group contains fully reconstructed $B^0 \to D_s^-\pi^+$, $B^0 - arrow\ D^-\rho^+$, $B^0 \to D^{*-}\pi^+$ and $B^0 \to D_s^{*-}\pi^+$ decays (shown in purple in the figure). Finally, the fifth group contains combinatorial background, and is shown in blue in Fig. 6.3.

The results of the simultaneous mass fit to $B_s^0 \to D_s^\mp K^\pm$ candidates and backgrounds are shown in Fig. 6.4. The types of backgrounds that contribute to the fit to $B_s^0 \to D_s^\mp K^\pm$ candidates have the same topology as the ones that contribute to the $B_s^0 \to D_s^-\pi^+$ fit. In this case, an important contribution to background comes from the partially reconstructed decays $B_s^0 \to D_s^\mp K^{*\pm}$, $B_s^0 \to D_s^{*\mp}K^\pm$ and $B_s^0 \to D_s^{*\mp}K^{*\pm}$. These backgrounds are represented in light blue in Fig. 6.4. Partially reconstructed decays of the type $B_s \to D_s^{(*)}(\pi, \rho)$ are shown in blue in the figure. Misidentified decays of the Λ_b^0 such as $\Lambda_b^0 \to D_s^{(*)-}p$ and $\Lambda_b^0 \to \Lambda_c^+K^-$ are depicted

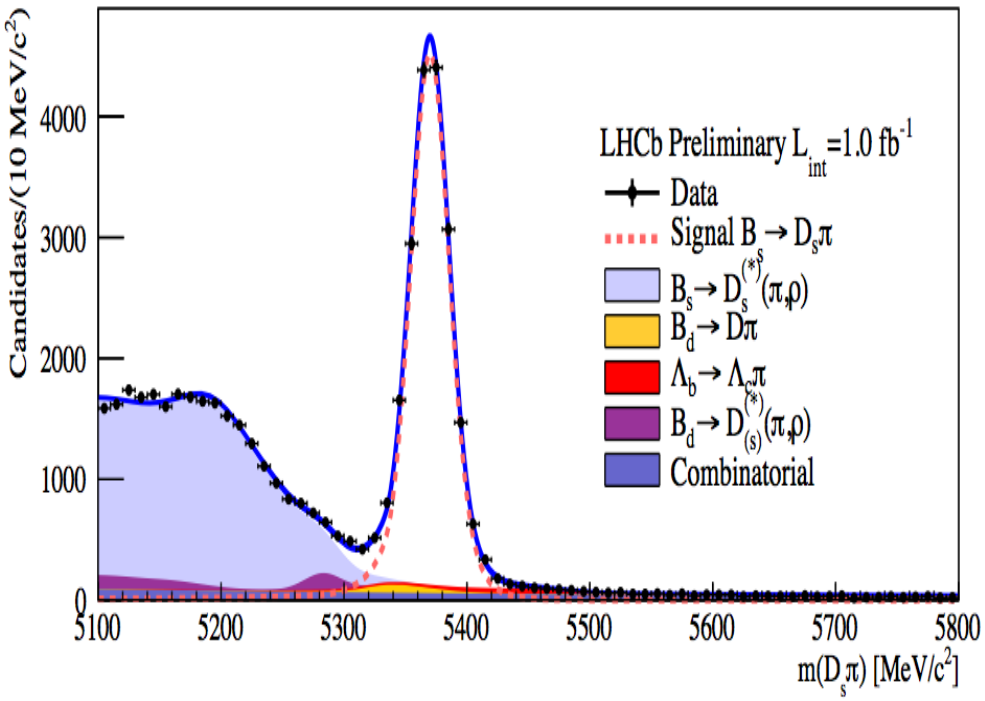

Figure 6.3: *Result of a simultaneous mass fit to the $B_s^0 \to D_s^- \pi^+$ candidates and all the contributing backgrounds [29].*

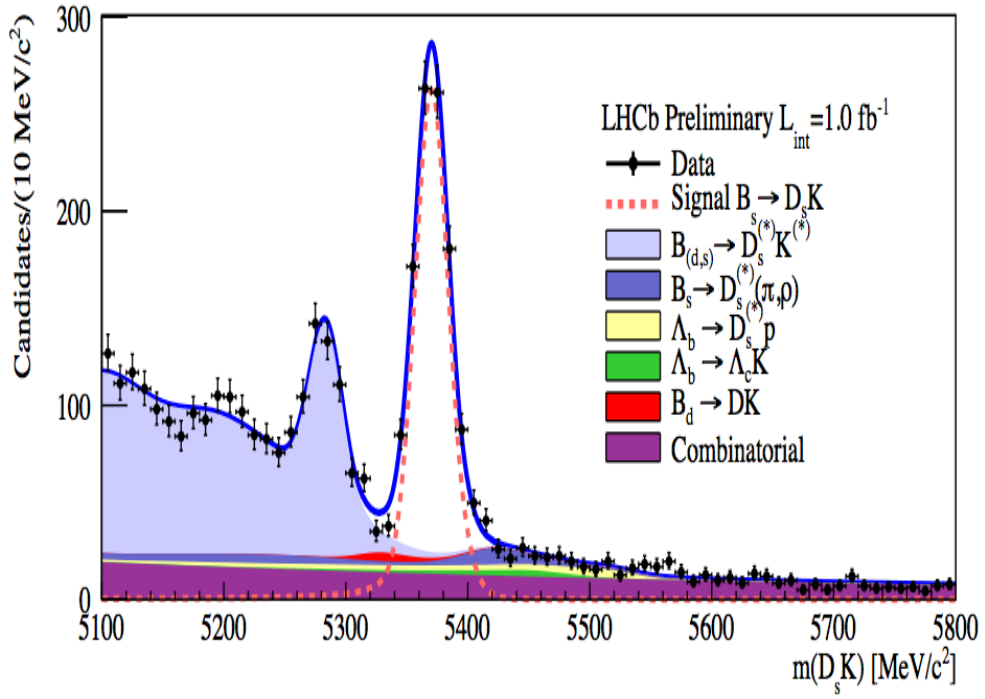

Figure 6.4: *Result of the simultaneous mass fit to the $B_s^0 \to D_s^\mp K^\pm$ candidates and all the contributing backgrounds [29].*

in yellow and green, respectively. In red, the distribution of the decay $B_d \to D^- K^+$ is shown, while the contribution of the combinatorial background can be seen in purple in the figure.

6.2 Fit Strategy

The fit strategy that was followed in this analysis is based on the *sPlot* technique [61, 62]. In this method, sWeights are assigned to all events in the data sample such that the sum of the sWeights for non-signal events is zero over the entire mass spectrum, and the sum of the sWeights over signal events is equal to the total signal yield. This assignment cancels the effect of any non-signal events in the sample. The sPlot technique and the way in which these sWeights are assigned will be discussed in more detail in Sec. 6.5.

By applying sWeights to our data sample, effectively a pure signal $B_s \to D_s\pi$ data sample is created. To this sample we fit the $B_s \to D_s\pi$ decay time probability density function (pdf) described by the decay rate Eq. 1.29-1.32. This pdf is convoluted with a proper-time resolution model, to account for uncertainties in the reconstructed proper-time of the B_s mesons. In addition, it is multiplied by an acceptance function, which represents the time-dependent efficiency of the detector and trigger and event selection. The signal pdf is then given by:

$$PDF_{Sig}(t, q_t, q_f) = a_{acc}(t) \cdot \int \frac{d\Gamma_{B_s \to D_s\pi}(t')}{dt'} \cdot R(t - t', \sigma_{t-t'}) dt' \tag{6.1}$$

or equivalently:

$$PDF_{Sig}(t, q_t, q_f) = a_{acc}(t) \ \cdot \ \frac{d\Gamma_{B_s \to D_s\pi}(t')}{dt'} \ \otimes \ R(t - t', \sigma_{t-t'}) \tag{6.2}$$

Here, $R(t - t', \sigma_{t-t'})$ is the decay time resolution model, $a_{acc}(t)$ the acceptance function, q_t is the flavor of the B_s^0 at production and q_f is the charge of the bachelor particle.

The time-dependent decay rate $d\Gamma(t')/dt'$ given in Eq. 1.29-1.32 is written assuming that the true flavour of the B_s meson at production is known. The initial flavour is assigned to the B_s candidate in a reconstruction process called flavour tagging. Since the assigned flavour is not always correct a mistag probability, that is predicted by the LHCb flavour tagging algorithms, is also assigned to each reconstructed. The algorithms used for flavour tagging will be discussed in Sec. 6.6, together with their performance for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays.

If we consider a non-zero mistag probability ω, the total number of tagged $B_s^0 \to D_s^- \pi^+$ candidates in the sample includes $B_s^0 \to D_s^- \pi^+$ decays in which the B_s^0 meson was tagged correctly as well as oscillated $\bar{B}_s^0 \to B_s^0 \to D_s^- \pi^+$ decays in which the $\bar{B}_s^0$ was tagged incorrectly as a B_s^0. For a given mistag probability ω the tagged pdf is:

$$\begin{aligned} PDF_{Sig}(t, q_t, q_f|\omega) = &\left((1-\omega) \frac{d\Gamma}{dt'} \left(B_s^0 \to D_s^- \pi^+ \right) + \omega \frac{d\Gamma}{dt'} \left(\bar{B}_s^0 \to B_s^0 \to D_s^- \pi^+ \right) \right) \\ &\otimes R\left(t - t', \sigma_{t-t'}\right) \ \cdot \ a_{acc}(t) \end{aligned} \tag{6.3}$$

where:

$$\begin{aligned} (1-\omega) \frac{d\Gamma}{dt'} \left(B_s^0 \to D_s^- \pi^+ \right) + \omega \frac{d\Gamma}{dt'} \left(\bar{B}_s^0 \to B_s^0 \to D_s^- \pi^+ \right) = \\ = e^{-t'/\tau} \left[\cosh\left(\frac{\Delta\Gamma_s t'}{2} \right) + (1 - 2\omega) \cos\left(\Delta m_s t' \right) \right] \end{aligned} \tag{6.4}$$

The decay time resolution model, $R(t-t', \sigma_{t-t'})$, consists in a triple Gaussian model while for the acceptance function, $a_{acc}(t)$, a power law parametrization was used. The decay time resolution model and acceptance function that are used in the fit will be discussed in detail in Secs. 6.3 and 6.4, respectively.

Given the mistag probability, ω, this pdf depends on the time t, the flavour tag of the B_s meson at production, q_t, and the charge of the bachelor particle, q_f. In order to obtain the expression of the final pdf that is used in the fit, the signal pdf in Eq. 6.2 must be multiplied with the distribution of the mistag probability, as:

$$PDF_{B_s \to D_s \pi} = PDF_{Sig}\left(t, q_t, q_f | \omega\right) \cdot P(\omega) \tag{6.5}$$

Given the signal pdf, the likelihood function of all events reads:

$$\mathcal{L} = \prod_{i=1}^{N} \left(PDF_{B_s \to D_s \pi}\left(t, q_t, q_f, \omega\right)\right) \tag{6.6}$$

where N is the total number of events in the data sample. To determine the physics parameters of the signal pdf the maximum-likelihood method is used. This method consists in selecting a set of values of the parameters of our model in a way such that the likelihood function is maximized. In other words, what we do is to maximize the probability of the observed data to be described by the function $PDF_{B_s \to D_s \pi}\left(t, q_t, q_f, \omega\right)$. In practice, it is more convenient to work with the logarithm of the likelihood function, or log-likelihood:

$$-log\mathcal{L} = -\sum_{i=1}^{N} log\left(PDF_{B_s \to D_s \pi}\left(t, q_t, q_f, \omega\right)\right) \tag{6.7}$$

This is because the maximum of a function is determined by taking the derivative of that function and then solving for the parameter that we want to maximize, and the logarithm is numerically easier to maximize compared to the likelihood function itself. Since the logarithm is a monotonically increasing function, its maximum will correspond to the maximum of the original likelihood. Since in our fit we look at the negative log-likelihood, to maximize the likelihood function we determine the minimum of $-log\left(\mathcal{L}\right)$.

6.3 Proper time resolution

Since a bias or mis-modeling of the proper time resolution affects the decay time pdf and thus also the fitted CP parameters, it is important for time resolution effects to be modeled correctly. The time resolution is implemented either by assuming an average resolution model or by assuming a model that has a different resolution for each event, by using the per-event decay time error determined in the reconstruction. The second approach is generally more precise but, since the per-event option is much more expensive in terms of computation time, a fixed resolution model was used in this analysis. A cross check showed that no bias in the fitted parameters occurs.

Two resolution models were studied; a double and a triple Gaussian of the form:

$$R_1(t - t', \sigma_1, \sigma_2) = f \cdot G(t - t'; \mu_1, \sigma_1) + (1 - f) \cdot G\left(t - t'; \mu_2, \sigma_2\right) \tag{6.8}$$

and

$$\begin{aligned} R_2(t - t', \sigma_1, \sigma_2, \sigma_3) = f_1 \cdot G(t - t'; \mu_1, \sigma_1) + f_2 \cdot G\left(t - t'; \mu_2, \sigma_2\right) \\ + \left(1 - f_1 - f_2\right) \cdot G\left(t - t'; \mu_3, \sigma_3\right) \end{aligned} \tag{6.9}$$

where:

$$G\left(t - t'; \mu, \sigma\right) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{6.10}$$

and all Gaussians have the same mean parameter ($\mu_1 = \mu_2 = \mu_3$). The parameters of the two models were determined by fitting the $(t_{rec} - t_{true})$ distribution, and the resulting parameters for the double and triple Gaussians are listed in Table 6.4.

Model	Parameter	Value
double Gaussian	σ_1	28.14 ± 0.58 fs
	σ_2	52.36 ± 1.03 fs
	μ	$(0.07 \pm 1.7) \cdot 10^{-4})$
	f	0.577 ± 0.028
triple Gaussian	σ_1	29.48 ± 0.027 fs
	σ_2	58.64 ± 0.063 fs
	σ_3	181.7 ± 4.9 fs
	μ	1.49 ± 0.14 fs
	f_1	0.595 ± 0.011
	f_2	0.386 ± 0.011

Table 6.4: *Parameters of the double Gaussian resolution model used for the $B_s^0 \to D_s^- \pi^+$ signal MC sample and triple Gaussian used for $B_s^0 \to D_s^- \pi^+$ data and $B_s^0 \to D_s^\mp \pi^\pm$ signal MC sample.*

The double Gaussian resolution model was found to accurately describe the time resolution of the $B_s^0 \to D_s^- \pi^+$ MC sample, as can be seen on the left of Fig. 6.5. For this reason it is used for $B_s^0 \to D_s^- \pi^+$ MC based studies, in which events are selected using a slightly different selection chain with respect to the one used on data.

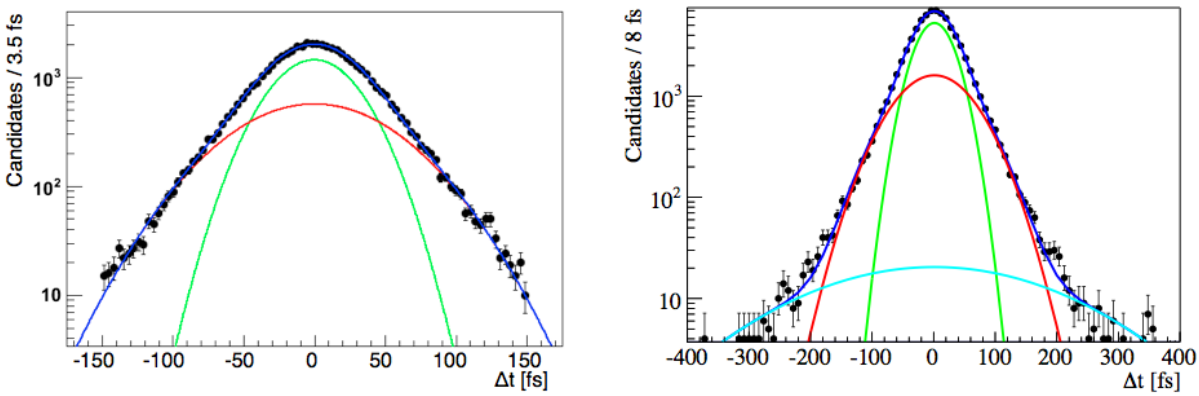

Figure 6.5: *Left: double Gaussian resolution model determined from the $B_s^0 \to D_s^- \pi^+$ signal MC sample. Right: triple Gaussian resolution model determined from the $B_s^0 \to D_s^\mp K^\pm$ signal MC sample.*

However, a triple Gaussian model was used in the tagging studies on $B_s^0 \to D_s^- \pi^+$ data, and was implemented in the proper time fit on real data. The same triple Gaussian model was also used to describe the time resolution in the physics fit on $B_s^0 \to D_s^\mp K^\pm$ events. The time resolution using a triple Gauss fit for the $B_s^0 \to D_s^\mp K^\pm$ signal MC sample is depicted on the right of Fig. 6.5 [29].

Since we are using the time resolution parameters from MC, we must take into account that there can be a difference in resolution models between data and MC, since for MC the

detector description during generation and reconstruction is identical while for data this is not guaranteed, as e.g. subtle remaining misalignments can have an additional contribution.

The decay $B_s^0 \to D_s^- \pi^+$ is used to determine a scale factor is obtained between the per-event time error estimated by the lifetime fit, and the true per-event time error. The two distributions that are obtained for the per-event time error are close in width to one, and the scale factor of the resolution model is given by the ratio of the widths of these two distributions, and account for the differences between data and MC. The method used to determine the scale factor of the resolution model is described in more detail in [60]. The scale factor used in this analysis is equal to 1.15. Therefore, the triple Gaussian with the widths scaled up by a factor 1.15 is used on data.

6.4 Time acceptance

In order to take into account the finite time acceptance of the trigger and event selection algorithms, the signal decay time pdf is multiplied by an acceptance function (see Eq. 6.2). The shape of the acceptance function was modeled using MC [29] from a sample of $B_s^0 \to D_s^- \pi^+$ events.

Various parametrizations of the decay time acceptance were tested, and it was found that the following power law parametrization best describes the acceptance:

$$a_{acc}(t) = \begin{cases} 0 & \text{if } (at)^n - b < 0 \text{ or } t < 0.2 \text{ ps} \\ (1 - \dfrac{1}{(1+(at)^n - b)}) \cdot (1 - \beta t) & \text{otherwise} \end{cases}$$

In this expression the parameter a, known as 'turn-on', models the steepness of the function. The exponent n and the offset b model the position of the turn-on, while the parameter β is needed to model the behaviour in time acceptance observed for large B_s meson decay times. In Fig. 6.6, the fit to the simulated MC sample of the proper time distribution is shown.

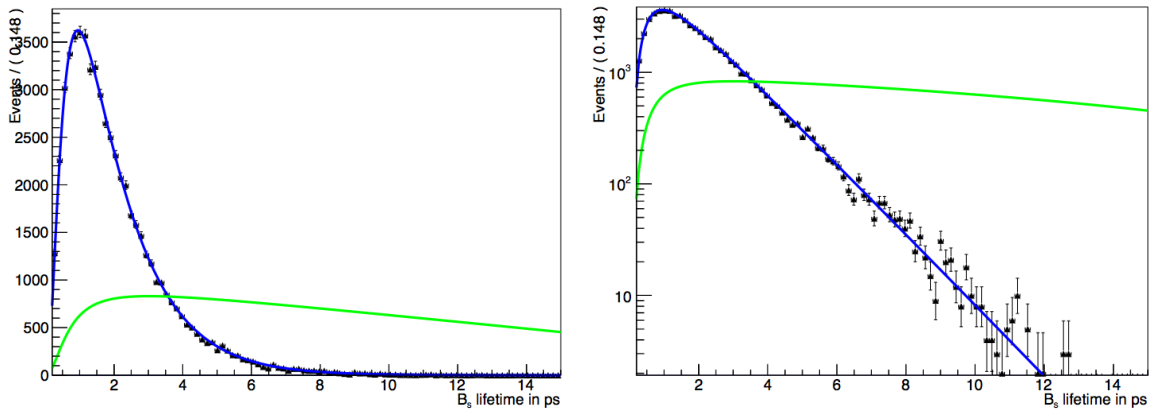

Figure 6.6: *Proper time distribution fitted to untagged $B_s^0 \to D_s^- \pi^+$ + cc in the simulated MC sample, in linear (left) and logarithmic (right) scale [29]. The data and the blue curve are the overall PDF; the green is the fitted acceptance function.*

The acceptance parameters are determined from a fit to $B_s^0 \to D_s^- \pi^+$ data, in which all events are treated untagged. In practice this means that both charge conjugations are summed and the physics decay rate simplifies to $\propto e^{-\frac{t}{\tau}} \cdot \cosh\left(\dfrac{\Delta\Gamma_s t}{2}\right)$ without oscillations, facilitating the acceptance determination. The resulting parameters are shown in Table 6.5.

Parameter	Value
β	0.0389 ± 0.0049 ps^{-1}
n	1.8882 ± 0.0802
b	0.0295 ± 0.0139
a	1.2213 ± 0.0419 ps^{-1}

Table 6.5: *Acceptance parameters determined from a fit to untagged* $B_s^0 \to D_s^- \pi^+$ *and charge conjugated data.*

6.5 Background subtraction using sWeights

If the shape of both signal and background events is known in a given variable (e.g. the B_s mass), the sFit method can be used to subtract background events in the signal region. Here, a short description of this method is given. Further details on the sFit method and on the *sPlot* technique can be found in [61] and [62].

In the sFit method, a weight (known as sWeight) is assigned to each event in the data sample in a way such that the signal component of an event can be isolated. The isolation of the signal component is based on a selected variable and works in a way that is analogous to the sideband subtraction method, where sWeights can be assigned to events in a data sample based on the B_s^0 mass distribution. In a likelihood fit, once the signal component of the mass distribution has been isolated, only the signal pdf will be left when the time fit is performed.

The sWeights are obtained such that, by construction, the sum over all sWeights gives the total signal yield, N_{sig}. Also by construction, the sum of the sWeights over background-only events will be equal to zero. Therefore, when the decay-time fit is performed, the background events drop because their weights average out, and only the signal events are left. For a sample of weighted candidates with weights w_i, we must make sure that a candidate contributes to the likelihood proportional to its weight, $\mathcal{L} = -\sum w_i(P(t_i, ...))$. This is analogous to filling a histogram with weighted events, e.g. in a sideband subtraction method.

The signal and background components can be modeled in a given variable which is used to calculate the sWeights (here the reconstructed B_s mass), and these sWeights can then be applied to other variables (such as time or mistag probability). The convenience of using the sFit method lies in the fact that it works for arbitrary shapes of the signal and background distributions, as long as the discriminating and signal variables are independent. The likelihood function obtained by applying the event weights results in a signal distribution which is free of any background component on a statistical level. This means, however, that the statistical errors on the parameters extracted from the fit might be larger than those resulting from fits in which other methods of background modeling are employed. However, the fact that no simulation for background modeling is needed might reduce the systematic uncertainty on the extracted parameters, since it reduces the systematic uncertainty introduced by an incomplete modeling of the background.

6.6 Tagging Performance

In $B_s^0 \to D_s^{\mp} K^{\pm}$ decays the time dependent CP violation observables are proportional to the amplitude of the observed decay rate oscillations, requiring the flavour of the B_s meson at production to be known. Although non-perfect flavour tagging procedure leads to a reduced sensitivity, an unbiased determination is possible by correcting for the wrong flavour tag probability. However, care must be taken since a bias in the assumed flavour tag fraction directly leads to a bias in the CP violation observable. In chapter 7 we present a tagging calibration method specifically developed for $B_s \to D_s h$ events.

In the flavour tagging procedure the initial flavour of the B_s meson is identified by determining if it contains a b or a $\bar{b}$ quark. There are two groups of algorithms that are used for flavour tagging: opposite-side (OS) tagging algorithms and same-side (SS) tagging algorithms.

At LHCb, b-quarks are produced in pairs ($b\bar{b}$). Opposite-side tagging algorithms determine the flavour of a given B meson (signal B) by exploiting the flavour of the other B hadron (tagging B) that was produced from the $b\bar{b}$ pair [63]. In particular, OS tagging algorithms identify either the charge of the lepton originating from semileptonic b-decays (this can be an electron or a muon), the charge of the kaon resulting from the $b \to c \to s$ transition, or the charge of the secondary vertex that is reconstructed from the products of b-hadron decays. Same-side tagging algorithms instead look at the correlation of the signal B with particles produced in the hadronization process [64]. To create a B_s meson an s quark is needed. Strange quarks are not contained in protons. Therefore, to get an s-quark an $s\bar{s}$ pair must be taken from the vacuum. While the s quark of the pair then combines with a $\bar{b}$-quark to form a B_s meson the $\bar{s}$, which is close in phase space to the B_s, might form a positively charged kaon with a u-quark from somewhere. The same argument holds for a $\bar{B}_s$ with all charges reversed. By determining the charge of the kaon it is then possible to determine wether it was an s or an $\bar{s}$ to form the B_s meson, thus determining the flavour of the B_s.

As the tagging decision is not always correct, the measured CP asymmetry is reduced proportional to the effective tagging power ϵ_{eff}, given by:

$$\epsilon_{eff} = \epsilon_{tag} D^2 = \epsilon_{tag}(1 - 2\omega)^2 \tag{6.11}$$

where ϵ_{tag} is the tagging efficiency and D the dilution, which depends on the mistag fraction ω; the fraction of wrongly tagged B_s candidates. In case of perfect tagging the mistag fraction ω is zero and the dilution is equal to one. Instead, in case the tagging is completely random (ω=50%) the dilution is equal to zero.

From the tagging reconstruction algorithm a tag (B_s^0 or $\bar{B}_s^0$) together with a mistag probability η are obtained. The reconstructed mistag η must be calibrated in order to obtain an unbiased mistag value of the real ω, the OS tagging algorithms are calibrated on data from the self-tagging decays $B^+ \to J/\psi K^+$ [65]. In these self-tagging decays a direct comparison between the tagging algorithm and the signal B decay products can be made, allowing a data driven calibration. Since the calibration for the SS taggers has not been done on this benchmark calibration, also in the analysis in this thesis only the information from the OS tagging was used, in order to make a comparison of the two methods.

In this thesis, the tagging calibration is performed on $B_s^0 \to D_s^- \pi^+$ data and then applied to $B_s^0 \to D_s^\mp K^\pm$ data, since these two decays modes have a similar topology. The $B_s^0 \to D_s^- \pi^+$ is used as a control channel, since in this decay the flavour of the B_s^0 meson is unambiguously determined by the charge of the final state pion making it a self-tagging decay. The observed decay rate oscillations therefore have a maximal amplitude as explained in Sec. 1.6. However, to perform the tagging calibration on $B_s^0 \to D_s^- \pi^+$ data and then apply it to $B_s^0 \to D_s^\mp K^\pm$ data, the tagging algorithms must perform exactly the same on the two channels. In Fig. 6.7, the distribution of the reconstructed mistag probability η for MC-truth matched simulated signal candidates of the two channels are compared. As can be seen from the figure, there is good agreement between the performance of the taggers in the two channels.

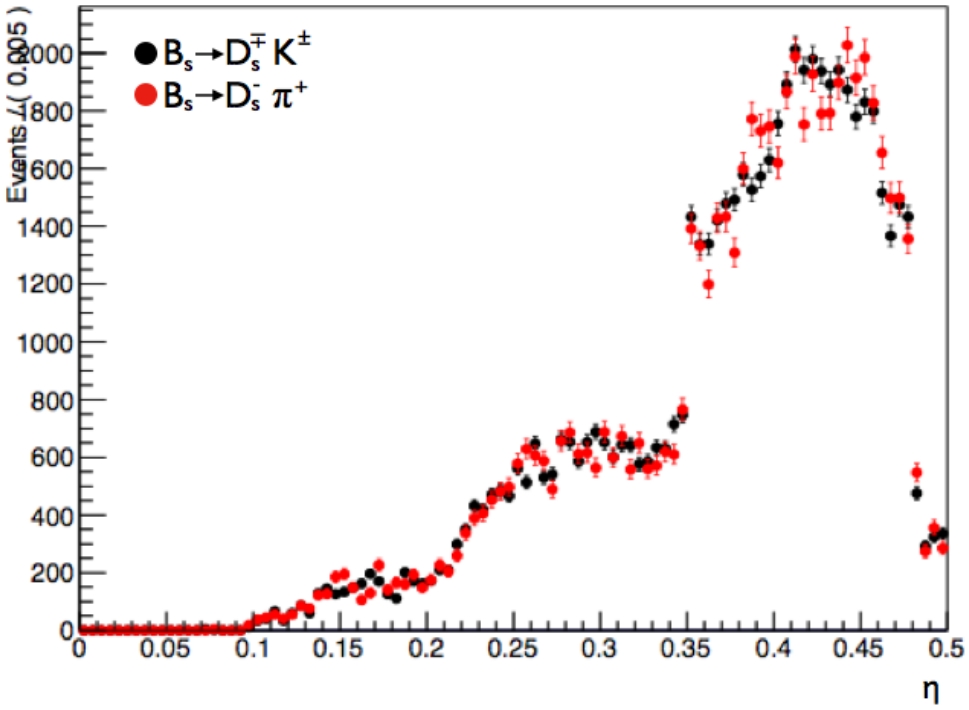

Figure 6.7: *Distribution of the reconstructed mistag probability η for the OS combination tagger in simulated data. Black is* $B_s^0 \to D_s^{\mp} K^{\pm}$ *and red is* $B_s^0 \to D_s^{-} \pi^{+}$ *[29].*

Chapter 7

Calibration of the mistag probability η on $B_s^0 \to D_s^- \pi^+$ data

To correctly extract the CP violation parameter γ from the data of $B_s^0 \to D_s^\mp K^\pm$ events the fraction of incorrectly flavour tagged events must be known. In this chapter, the calibration of the assigned mistag probability η to reproduce the true mistag fraction ω is done using self-tagging $B_s^0 \to D_s^- \pi^+$ events. The method relies on the fact that the decay time dependent flavour oscillations for self-tagging decays have a known, maximal amplitude.

The standard LHCb calibration of the mistag probability η is performed using the benchmark channel $B^+ \to J/\psi K^+$. Since the tagging quality might depend on the kinematics of the final state particles we chose here to apply a calibration with $B_s^0 \to D_s^- \pi^+$ events for the $B_s^0 \to D_s^\mp K^\pm$ channel, since the final states of these decays are very similar. It is, however, useful to compare the calibration output of the two methods.

In Sec. 7.1 the calibration procedure using $B_s^0 \to D_s^- \pi^+$ events is validated using toy MC studies, in Sec. 7.2 the method is tested on a full `GEANT` MC data sample and in Sec. 7.3 the calibration is performed on the 2011 data sample of $B_s^0 \to D_s^- \pi^+$ events. The systematic errors of the procedure are discussed in Sec. 7.4.

7.1 Calibration of mistag fraction using mistag categories on toy MC

We generate a toy MC data sample with 26,000 $B_s^0 \to D_s^- \pi^+$ signal events and their charge conjugates. The $B_s^0 \to D_s^- \pi^+$ decay has identical topology as the $B_s^0 \to D_s^\mp K^\pm$ decay, apart from a slightly different momentum of the bachelor pion and kaon which is ignored here. Since the decay $B_s^0 \to D_s^- \pi^+$ is flavour specific, the flavour of the decaying B_s^0 meson can be determined by measuring the charge of the bachelor π. The events are generated with an event-by-event value of the true mistag probability ω. The per-event mistag distribution generated is depicted in Fig. 7.1, and is constructed to roughly reproduce the shape of the mistag in real data (see Fig. 6.7).

As shown in the figure, the mistag fraction is taken to increase quadratically until a certain value ω_c, then to fall linearly for values higher than ω_c. The value of ω_c is chosen such that the average mistag is equal to 36.96%. In the three regions, $P(\omega)$ is described by the functions:

$$0 \qquad if \quad \omega \leq \omega_0 \tag{7.1}$$

$$\frac{(\omega - \omega_0)^2}{(\omega_c - \omega_0)^2} \qquad if \quad \omega_0 \leq \omega \leq \omega_c \tag{7.2}$$

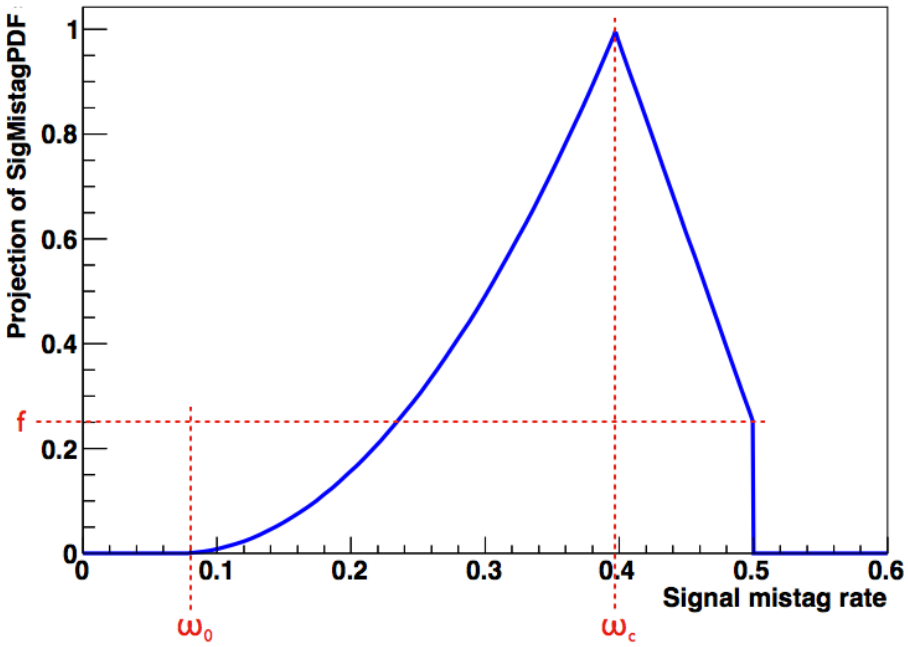

Figure 7.1: *Per-event mistag distribution.*

and

$$\left(1 - \frac{(1-f)(\omega - \omega_c)}{(0.5 - \omega_c)}\right) \qquad if \quad \omega_c \leq \omega \leq 0.5 \tag{7.3}$$

The above description gives a per-event mistag distribution which roughly corresponds to the shape of the distribution that is observed in data, and is convenient for two different reasons. First, when described by these functions, this distribution used together with the decay rate PDF of $B_s^0 \to D_s^- \pi^+$, as described in Sec. 6.2, can be integrated analytically, which considerably increases the computational speed of the generation and fitting of decay rate pdfs. Second, the value of the average mistag probability can be easily chosen order to correspond to the value observed in data. This is done by tuning the parameters ω_c, ω_0 and f.

For the validation of the method, we put the reconstructed mistag η equal to the true mistag ω in the toys and try to reproduce the perfect calibration using the parametrization:

$$\omega = p_0 + p_1(\eta - \langle\eta\rangle)$$

where we should find: p_0=$\langle\eta\rangle$ and p_1=1.

In order to calibrate the mistag fraction η, we divide our data into categories in which events are classified according to the reconstructed mistag probability in that event. We select these categories over the whole range $\eta \in [0, 0.5]$, such that all categories contain a comparable number of events.

The mistag categories that were selected are listed in Table 7.1, where the number of events having a value of η that falls into a selected range is listed.

We subsequently fit for the mistag fraction ω for each category using the pdf as given in Eq. 1.29-1.32 and making use of the self-tagging character of the decays. This procedure is repeated for 500 toy experiments of 26,000 events each. The average value of the reconstructed mistag $\langle\eta_i\rangle$ and the fitted value of $\langle\omega_i\rangle$ for each category i are also in the table. The distribution of ω for categories 0 and 4 are shown as examples in Fig. 7.2 (left) together with the pull distributions for these mistag fractions (right), where the pull variable is defined as $(\omega - \omega_{true})/\sigma_\omega$

Mistag Category	η-Range	Number of events	$\langle\eta_i\rangle$	$\langle\omega_i\rangle$
0	$[0-0.25]$	2485	0.205	0.205 ± 0.001
1	$[0.25-0.32]$	3994	0.289	0.289 ± 0.001
2	$[0.32-0.36]$	3606	0.341	0.342 ± 0.001
3	$[0.36-0.39]$	3480	0.375	0.375 ± 0.001
4	$[0.39-0.42]$	4275	0.405	0.406 ± 0.001
5	$[0.42-0.45]$	4195	0.434	0.433 ± 0.001
6	$[0.45-0.50]$	3965	0.471	0.471 ± 0.001

Table 7.1: *Selected mistag categories; the average mistag $\langle\eta_i\rangle$ and fitted mistag $\langle\omega_i\rangle$ for each category are listed.*

and ω_{true} is taken from the MC information. Similar distributions were observed for all other categories. From the pull distributions it is observed that the measured values of ω are in good agreement with the expected ones.

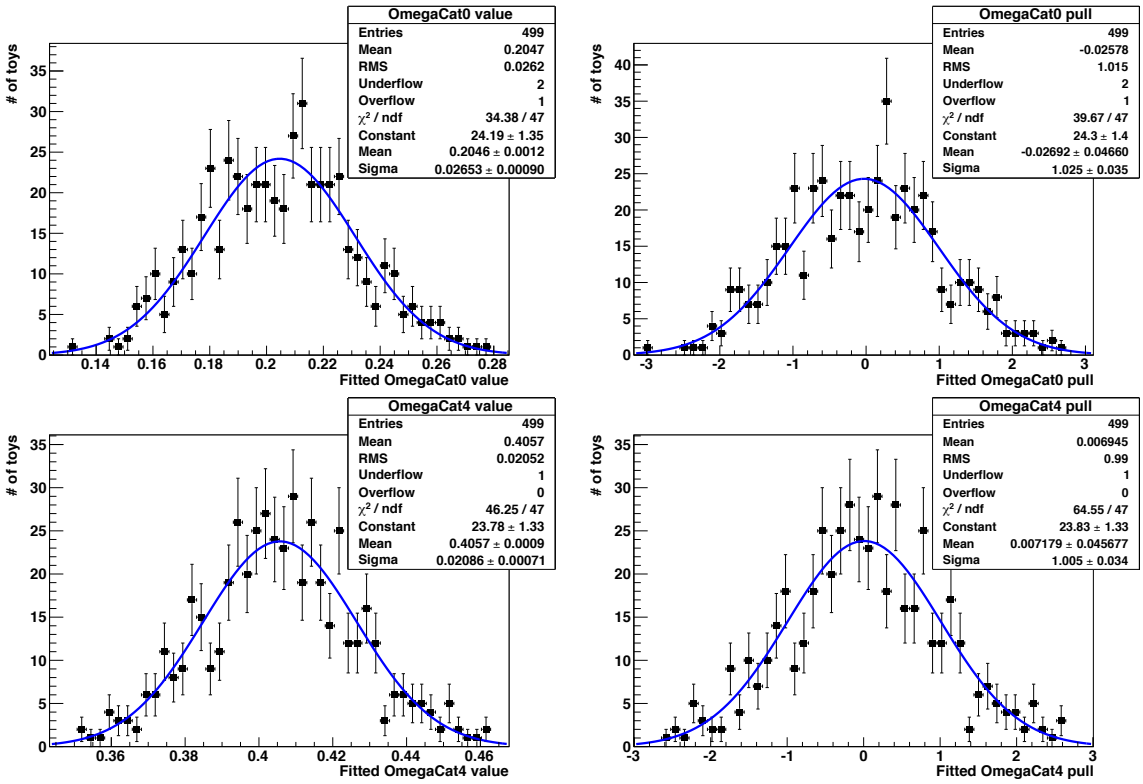

Figure 7.2: *Distribution of the mistag fraction ω (left) and respective pulls (right) for the selected categories 0 (top) and 4 (bottom).*

For each category, we calculate the reconstructed average mistag fraction, η, from the per-event mistag distribution. The value of η for each category is found by calculating:

$$\langle\eta_i\rangle = \frac{\int_{a_i}^{b_i} \eta' \cdot P(\eta')\ d\eta'}{\int_{a_i}^{b_i} P(\eta')\ d\eta'} \tag{7.4}$$

where a_i, b_i are the lower and upper bounds of the different categories i. Similarly, we integrate the per-event mistag distribution over the entire range $\eta \in [0, 0.5]$ and find:

$$\langle \eta \rangle = 0.3696$$

which indeed correctly corresponds to the assumed average mistag value.

To obtain the calibration curve, we look at the measured mistag fraction ω as a function of the reconstructed mistag η, for each category. The ω vs η distribution for one toy experiment is shown in Fig. 7.3. Similar distributions were observed for all toys.

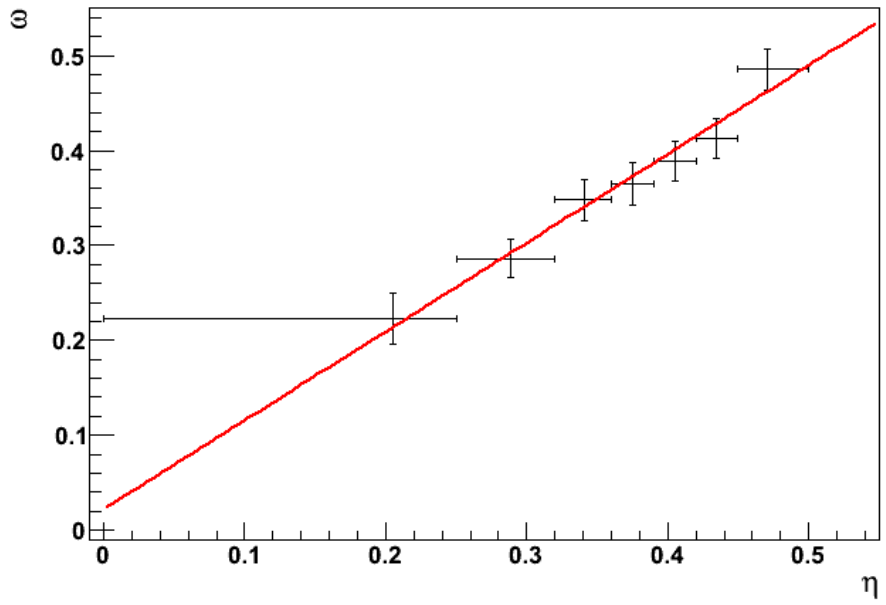

Figure 7.3: *Fitted mistag fraction ω as a function of the reconstructed mistag η, for a single toy experiment of 26,000 $B_s^0 \to D_s^- \pi^+$ events.*

Using the parametrization $\omega = p_0 + p_1(\eta - \langle \eta \rangle)$, we fit for the parameters p_0 and p_1. Since for these toys the calibration is assumed to be perfect, $\langle \omega_i \rangle$ should be equal to $\langle \eta_i \rangle$ for each mistag category i, and we expect to find $p_0 = \langle \eta \rangle$ and $p_1 = 1$. The distribution of the calibration parameters p_0, p_1 and their respective pulls, for 500 toy experiments, are shown in Fig. 7.4.

Indeed, the fitted values of p_0=0.3692 $\pm$ 0.0086 and p_1=1.000 $\pm$ 0.109 correspond to the expected input values, indicating that the self-tagging feature of $B_s^0 \to D_s^- \pi^+$ events can indeed be used to calibrate the wrong tag fraction.

7.2 Calibration of the mistag rate η on simulated $B_s^0 \to D_s^- \pi^+$ events

The fact that the results obtained with the toys are correct shows that the toy experiments are 'internally' consistent, since data is generated using the same pdf as the one that is used in the fitting. It is therefore important to test the method on fully simulated MC events, where the data is generated through the full chain of detector simulation and event selection.

The kinematical selection described in Sec. 6.1 is applied to the MC sample, resulting in a sample of 50552 $B_s^0 \to D_s^- \pi^+$ events. The proper time resolution is modeled with a double Gaussian distribution, as was described in Sec. 6.3.

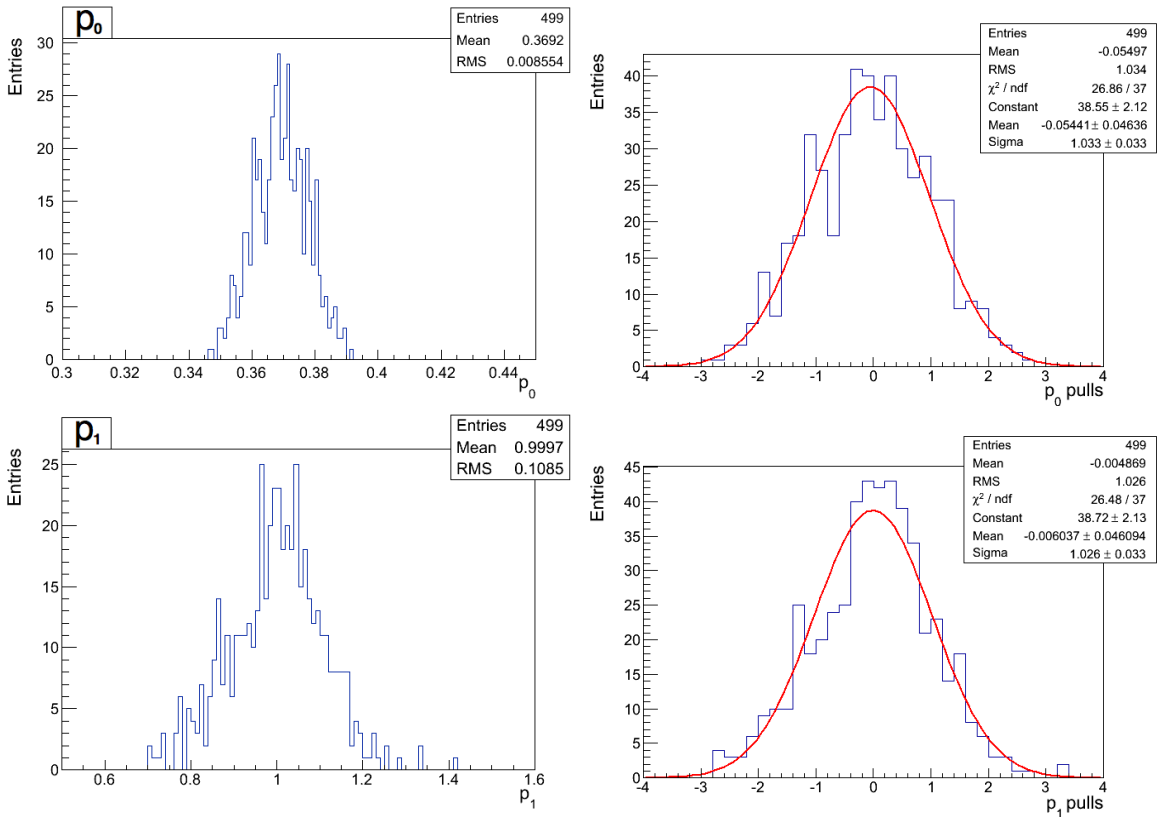

Figure 7.4: *Distribution of p_0 (top-left) and p_1 (bottom-left), and pull distributions for p_0 (top-right) and p_1 (bottom-right). The distributions are shown for 499 toy experiments, since for one toy experiment the fit did not converge.*

The MC data is divided into the same mistag categories as described in Sec. 7.1. In Table 7.2, the selected mistag categories and the number of events with a value of η that falls into the range of each category are shown. The MC truth mistag value ω_{true}, the average reconstructed mistag for each category $\langle \eta_i \rangle$, and the calibrated value of $\langle \omega \rangle$ resulting from the fit on MC data in the different categories are listed in the table.

Mistag Category	η-Range	Number of events	ω_{true}	$\langle \eta_i \rangle$	$\langle \omega_i \rangle$
0	$[0 - 0.25]$	1795	0.207	0.199	0.187 ± 0.020
1	$[0.25 - 0.32]$	2067	0.270	0.288	0.276 ± 0.019
2	$[0.32 - 0.36]$	2451	0.326	0.342	0.349 ± 0.018
3	$[0.36 - 0.39]$	2417	0.345	0.375	0.335 ± 0.018
4	$[0.39 - 0.42]$	2617	0.398	0.405	0.411 ± 0.018
5	$[0.42 - 0.45]$	2471	0.426	0.435	0.426 ± 0.018
6	$[0.45 - 0.50]$	1775	0.438	0.465	0.438 ± 0.022

Table 7.2: *Range of the selected mistag categories. For each category, the number of events, the MC-true mistag ω_{true}, the reconstructed average mistag $\langle \eta_i \rangle$ and the calibrated mistag $\langle \omega_i \rangle$ are listed.*

For each category, the true mistag probability ω_{true} is calculated by counting the number of

events in that category for which the flavour of the reconstructed B_s^0 is not equal to its 'true' flavour, divided by the number of tagged events. In the Monte Carlo the standard tagging algorithms are tuned to produce unbiased results such that the average reconstructed tag $\langle \eta_i \rangle$ is close to ω_{true}. The average mistag ω_{true} is equal to:

$$\langle \omega_{true} \rangle = 0.3485$$

while the average reconstructed mistag η is:

$$\langle \eta \rangle = 0.3629$$

The fitted mistag probability ω as a function of the reconstructed mistag η for the seven mistag categories is depicted in Fig. 7.5.

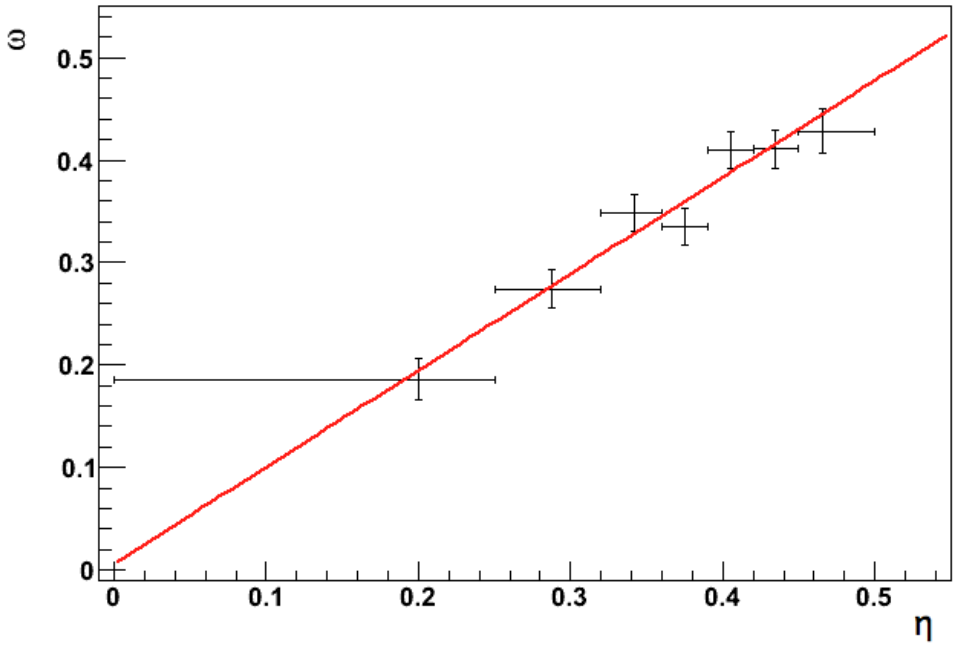

Figure 7.5: *Calibrated mistag ω as a function of the reconstructed mistag η, for the 7 selected mistag categories.*

The predicted mistag probability ω is then parametrized according to the function:

$$\omega = p_0 + p_1(\eta - \langle \eta \rangle) \tag{7.5}$$

The parameters resulting from the fit are:

$$p_0 = 0.3470 \pm 0.0071$$
$$p_1 = 0.943 \pm 0.088$$

showing that both parameters p_0 and p_1 are compatible within 2σ with the expected values of $\langle \eta \rangle$ and 1, given the respective uncertainties.

To show explicitely that the calibration gives correct results, in Fig. 7.6 the fitted mistag ω is shown as a function of ω_{true}.
Here, the calibrated mistag probability ω is then parametrized according to the function:

$$\omega = p_0 + p_1(\omega_{true} - \langle \omega_{true} \rangle) \tag{7.6}$$

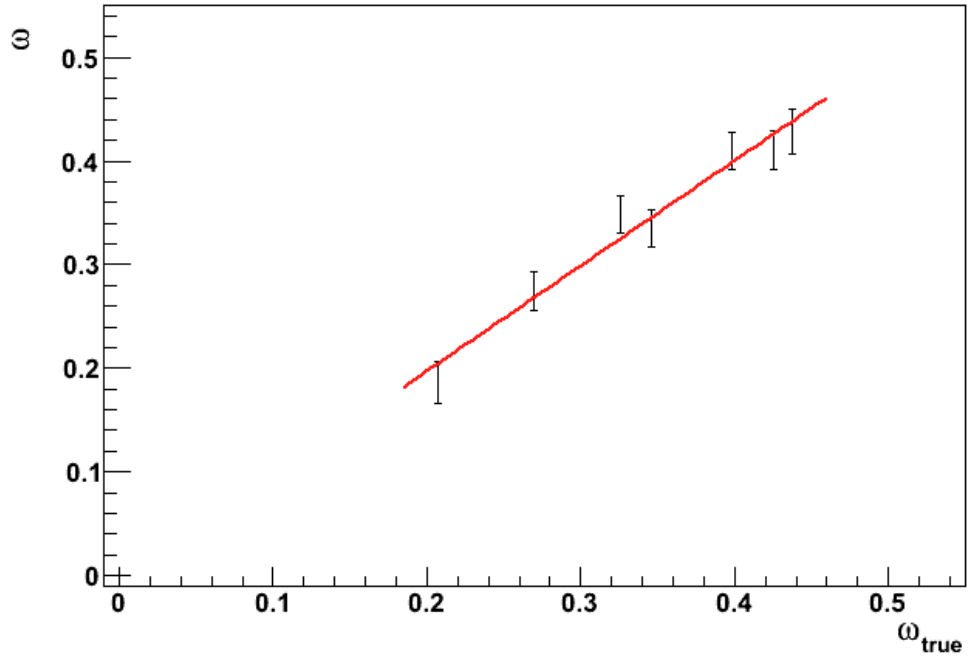

Figure 7.6: *Calibrated mistag ω as a function of the true mistag probability ω_{true}, for the 7 selected mistag categories.*

In case of MC truth, the parameters p_0 and p_1 should be equal to $\langle\omega_{true}\rangle$ and 1, respectively. The parameters resulting from the fit in the case ω_{true} is used are equal to:

$$\begin{aligned} p_0 &= 0.3471 \pm 0.0071 \\ p_1 &= 1.016 \pm 0.094 \quad , \end{aligned}$$

confirming $p_0 = \langle\omega_{true}\rangle$ and $p_1 = 1$.

The results demonstrate the stability and accuracy of the method used for the calibration of the mistag probability η. The same method is next employed to calibrate the reconstructed mistag η using the 2011 LHCb $B_s^0 \to D_s^- \pi^+$ real data sample.

7.3 Calibration of η on $B_s^0 \to D_s^- \pi^+$ data

Since our method used to calibrate the mistag probability is validated on simulated MC data, the calibration of η is finally performed, in a fit, on the 2011 LHCb $B_s^0 \to D_s^- \pi^+$ data sample, pre-selected with the stripping cuts described in Sec. 6.1. The input parameters of the fit and the results are described in this section, while the corresponding systematic uncertainties on the measurement of the parameters p_0 and p_1 is discussed in the following section.

The values to which the mean decay width Γ_s and the decay width difference $\Delta\Gamma_s$ are fixed have been measured using a sample of $B_s^0 \to J/\psi\ \phi$ events extracted from 1 fb^{-1} of pp collisions collected in 2011 [66], and are:

$$\Gamma_s = 0.6580 \pm 0.0054(\text{stat}) \pm 0.0066(\text{syst})\ \text{ps}^{-1}$$

and

$$\Delta\Gamma_s = 0.116 \pm 0.018(\text{stat}) \pm 0.006(\text{syst})\ \text{ps}^{-1}$$

The scale factor used to scale the calculated per-event error in the resolution model of the decay time (see Sec. 6.3), which is determined by comparing the pulls of the proper time error measured in data and MC [29], is set to:

$$\mathrm{SF} = 1.15^{+0.06}_{-0.10}$$

The acceptance parameters are determined using $B_s^0 \to D_s^- \pi^+$ data itself, treating all events as untagged. The resulting values of the acceptance parameters *β, n, a, b* were given in Table 6.5 and the function used to model the acceptance is defined in Sec 6.4.

The number of $B_s^0 \to D_s^- \pi^+$ signal events is extracted from the data sample, by constructing the sum of the sWeights over all events in the sample: $\sum sWeights = N_s$, where N_s corresponds the signal yield from the mass fit. A total of $N_s = 27965 \pm 196$ $B_s^0 \to D_s^- \pi^+$ events are present in the selected data sample.

The tagging efficiency ϵ_{tag} is calculated to be 39.27±0.52 %, and is fixed in the fit. The data was subdivided into the same mistag categories as before. In Table 7.3, the number of events in each category and the average value of η measured in that category are listed.

Mistag category	η-Range	Number of events	$\langle\eta_i\rangle$	$\langle\omega_i\rangle$
0	$[0 - 0.25]$	884.00	0.201	0.1828 ± 0.0401
1	$[0.25 - 0.32]$	1360.92	0.288	0.2249 ± 0.0326
2	$[0.32 - 0.36]$	981.75	0.342	0.3316 ± 0.0437
3	$[0.36 - 0.39]$	1375.03	0.376	0.3586 ± 0.0375
4	$[0.39 - 0.42]$	1772.26	0.406	0.4595 ± 0.0319
5	$[0.42 - 0.45]$	2067.93	0.435	0.4790 ± 0.0230
6	$[0.45 - 0.50]$	2540.96	0.476	0.5002 ± 0.0270

Table 7.3: *Range of the selected mistag categories. For each category, the number of events with η in the category's range, the average value $\langle\eta_i\rangle$, and the calibrated wrong tag $\langle\omega_i\rangle$ in that category are given.*

The fitted mistag probability $\langle\omega_i\rangle$ is then compared to the reconstructed mistag $\langle\eta_i\rangle$ in each category. In Fig. 7.7, the calibrated mistag $\langle\omega_i\rangle$ as a function of $\langle\eta_i\rangle$ is shown.

The mistag ω is again parametrized according to the function:

$$\omega = p_0 + p_1(\eta - \langle\eta\rangle)$$

where the average mistag probability $\langle\eta\rangle$, calculated over the entire mistag range $[0 - 0.5]$, is:

$$\langle\eta\rangle = 0.3833$$

The resulting parameters p_0 and p_1 are:

$$p_0 = 0.395 \pm 0.013$$
$$p_1 = 1.369 \pm 0.152$$

It can be seen that the value of p_0 is compatible with the value of the average mistag rate $\langle\eta\rangle$ within the statistical uncertainty. On the other hand, the value of p_1 is not compatible with the expected value of 1.

The nominal LHCb taggers are calibrated using the self-tagging decay channel $B^+ \to J/\psi K^+$, exploiting the high statistics present for this channel. The parameters p_0 and p_1 determined by the tagging group on $B^+ \to J/\psi K^+$ data [67] are:

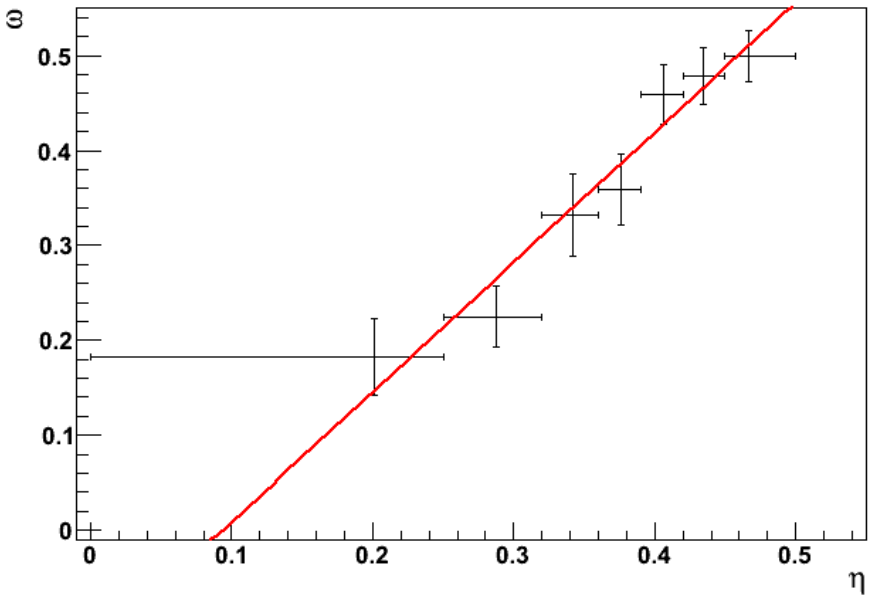

Figure 7.7: *Calibrated mistag ω as a function of the reconstructed mistag η, for the 7 selected mistag categories.*

$$p_0 = 0.392 \pm 0.002$$
$$p_1 = 1.035 \pm 0.021$$

The value of p_0 measured in our calibration is compatible with the result obtained by the tagging group, while there is a discrepancy between the value of p_1 obtained from our fit and the tagging group's result. A further investigation and the effect of the systematic uncertainties are discussed in detail in the following section.

7.4 Study of the systematics

In order to test the compatibility of our results for p_0 and p_1 with the results of the benchmark channel $B^+ \to J/\psi K^+$, the systematic uncertainty on our method is evaluated. Two independent systematic effects are evaluated:

- The dependence of the value of the externally fixed fit parameters. These are the PDF input parameters with an uncertainty that affects the measurement of p_0 and p_1, such as the mean decay width Γ_s, the decay width difference $\Delta\Gamma_s$, the scale factor of the proper time resolution model SF_{res}, the tagging efficiency ϵ_{tag}, the total signal yield N_s, the number of mistag categories N_{cats}, and the acceptance parameters β, n, a, b.

- The dependence due to the chosen sWeights assigned to all events. By applying sWeights to the events in the data sample the background yield is subtracted from the total number of events in the sample. Therefore, if the treatment of the background is not performed correctly the resulting tagging calibration will be imprecise. An alternative method will be used to test for this effect.

7.4.1 PDF input parameters

To see how a variation of the input parameters of the pdf affects the measurement of p_0 and p_1 the parameters, generally labeled x_i are varied within their uncertainty, and each time the calibration of η was performed and p_0 and p_1 extracted neglecting any correlations between the parameters. To do this, for each x_i the values of p_0 and p_1 are measured twice: once with $(x_i - \sigma_{x_i})$ and once with $(x_i + \sigma_{x_i})$ as input for the mistag calibration. The resulting values of p_0 and p_1 are indicated as $p_0^{\pm}$ and $p_1^{\pm}$. The values of $p_0^{\pm}$ and $p_1^{\pm}$ are then compared with the results of the nominal fit: $p_0 = 0.395 \pm 0.013$ and $p_1 = 1.369 \pm 0.152$. The effect that a variation in the number of tagging categories N_{cats} has on the measurement of p_0 and p_1 was studied by once using only 4 mistag categories (p_0^-, p_1^-) and once using 14 mistag categories (p_0^+, p_1^+).

The difference between the results of the nominal fit and the results obtained when varying the parameters x_i represents the systematic uncertainty induced on p_0 and p_1 in the different cases, and is shown in Table 7.4. In most cases the effect is negligible; only the scale factor of the proper time resolution model SF and the number of mistag categories N_{cats} contribute significantly to the overall systematic error on the calibration parameters p_0 and p_1.

	$x_i \pm \sigma_{x_i}$	$\lvert p_0 - p_0^+ \rvert$	$\lvert p_0 - p_0^- \rvert$	$\lvert p_1 - p_1^+ \rvert$	$\lvert p_1 - p_1^- \rvert$
Γ_s	0.658 ± 0.009	2.90×10^{-5}	2.99×10^{-5}	3.69×10^{-4}	3.80×10^{-4}
$\Delta\Gamma_s$	0.116 ± 0.019	2.99×10^{-8}	2.90×10^{-6}	4.86×10^{-3}	5.55×10^{-3}
SF_{res}	1.15 ± 0.06	3.59×10^{-3}	5.51×10^{-3}	4.37×10^{-2}	6.88×10^{-2}
N_s	27,965 ± 196	4×10^{-6}	4×10^{-6}	4×10^{-5}	4×10^{-5}
ϵ_{tag}	0.393 ± 0.005	4×10^{-6}	4×10^{-6}	4×10^{-5}	4×10^{-5}
N_{cats}	4, 7, 14	1.49×10^{-3}	4.22×10^{-4}	4.15×10^{-2}	2.33×10^{-2}
Acc. β	0.039 ± 0.005	0	1×10^{-6}	1×10^{-5}	0
Acc. n	1.888 ± 0.080	1×10^{-6}	1×10^{-6}	0	1×10^{-5}
Acc. b	0.029 ± 0.014	1×10^{-6}	1×10^{-6}	0	1×10^{-5}
Acc. a	1.221 ± 0.042	1×10^{-6}	1×10^{-6}	0	0

Table 7.4: *Variations on the calibration parameters p_0, p_1 determined when the input parameters x_i are varied within their statistical uncertainty.*

7.4.2 sWeights

The main systematic uncertainty is related to the fact that background events are statistically subtracted by assigning sWeights. As described in Sec. 6.5, an sWeight is assigned to each event, which represents the fraction of signal pdf present in that event. After weighting each event with the corresponding sWeight, the resulting data sample should effectively consist only of signal events.

To set the scale of a possible systematic effect on the mistag fraction due to the treatment of background events an unweighted fit is performed in the B_s^0 mass window, and the values obtained for the calibration parameters p_0 and p_1 in the *weighted* and *unweighted* fit are compared. The unweighted fit is performed by setting all sWeights equal to 1, therefore effectively assuming that no background is present in the selected mass window. The B_s^0 mass window in which the two fits are performed corresponds to the mass range [5320, 5420] GeV/c^2.

The values of p_0 and p_1 extracted from the unweighted fit are:

$$p_0 = 0.4067 \pm 0.0097$$
$$p_1 = 1.0513 \pm 0.1218$$

and the difference between these values and the values of p_0 and p_1 extracted from the nominal fit is equal to:

$$|p_0^{nom} - p_0| = 0.012$$
$$|p_1^{nom} - p_1| = 0.318$$

Compared to the statistical uncertainty on the calibration parameters calculated in the nominal fit:

$$\sigma(p_0^{nom}) = 0.013$$
$$\sigma(p_1^{nom}) = 0.152$$

it is clear that the systematic uncertainty introduced by the treatment of the background cannot be ignored a priori.

To see how the background fraction affects the systematic uncertainty on p_0 and p_1 in the unweighted fit, it is repeated in two additional cases: when the B_s^0 mass window is half the size of the original one, so [5345,5395] GeV/c^2, and when it has twice its size, so [5270,5470] GeV/c^2. The average mistag probability $\langle\eta\rangle$, the total signal yield N_{Sig} and the calibration parameters are listed in Table 7.5 for the three mass windows and for the sFit.

B_s^0 mass window	N_{Sig}	$\langle\eta\rangle$	p_0	p_1
[5345,5395] GeV/c^2	23804 ± 154	0.386	0.399 ± 0.011	1.179 ± 0.1337
[5320,5420] GeV/c^2	28873 ± 170	0.385	0.407 ± 0.010	1.051 ± 0.122
[5270,5470] GeV/c^2	33756 ± 184	0.386	0.409 ± 0.009	0.960 ± 0.113
sFit	27965 ± 196	0.383	0.395 ± 0.013	1.369 ± 0.152

Table 7.5: *Results of the unweighted fit performed in three different mass windows, and of the sFit: N_{sig} and $\langle\eta\rangle$ are shown, together with the calibration parameters p_0 and p_1.*

As expected, the total signal yield is smaller when a narrow mass window is selected. The average value of the reconstructed mistag probability, $\langle\eta\rangle$, does not change significantly when the original mass window is reduced or increased, which implies that the selected background events are kinematically similar to the signal events.

The results show that the tagging calibration does depend on the amount of background in the unweighted fit. For smaller mass windows, i.e. a relative small amount of background, the calibration is closer to the sWeighted case. In the next section, an alternative approach is followed: instead of ignoring the background, a specific background decay time component is added to the pdf.

7.4.3 Study of background

To verify that the large difference in the calibration parameters p_0 and p_1 observed between the sFit method and the unweighted fits in progressively narrower mass windows is really due to neglected backgrounds, an alternative fit is performed including the background events and their tagging behaviour explicitly in the PDF, instead of subtracting them using sWeights.

The model used to describe the signal invariant mass distribution corresponds to a double Gaussian with the two Gaussians centered at the B_s^0 mass, with widths σ_1 and σ_2. The background distribution instead is described by the sum of a Gaussian, accounting for partially reconstructed decays of Λ_b^0, and an exponential, describing the contribution of combinatorial background. In Fig. 7.8, the invariant mass distribution of $B_s^0 \to D_s^-\pi^+$ decays is shown in linear and logarithmic scale. In both figures, the signal component is depicted in red while the background component is shown in green. The parameters that define these distributions are listed for both the signal and background components.

To describe the tagging behaviour of the background component, the signal distribution is multiplied by an additional pdf which accounts for the non-oscillating component of the proper time distribution. Any small contribution from B_d decays that might be present will to first

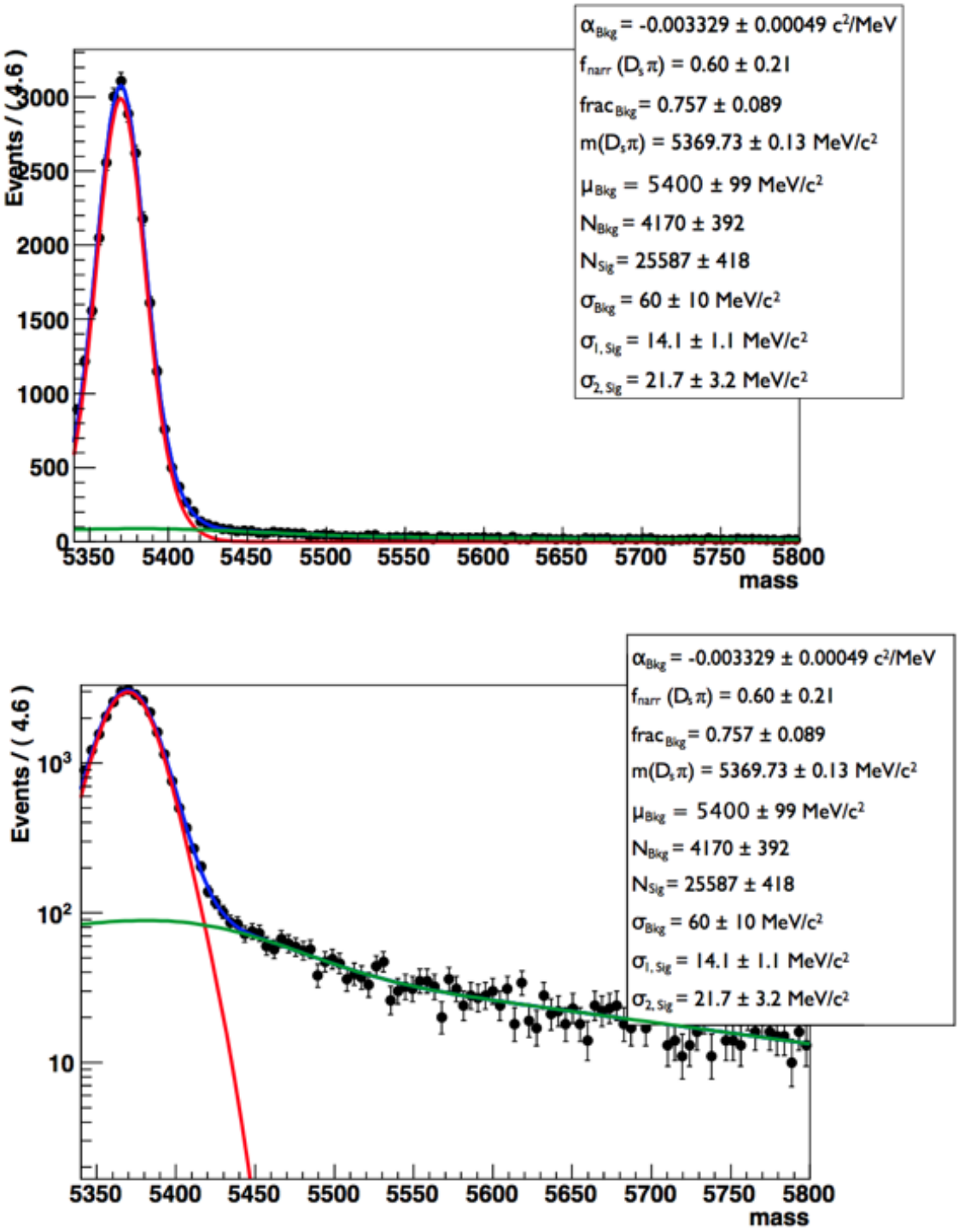

Figure 7.8: *Invariant mass distribution of $B_s^0 \to D_s^- \pi^+$ candidates, in linear scale (top) and logarithmic scale (bottom). The signal component is shown in red and the background component in green.*

order look like a non-oscillating background due to the relatively small B_d mixing rate when compared to the B_s. The proper time distribution of $B_s^0 \to D_s^- \pi^+$ decays is shown in linear and logarithmic scale in Fig. 7.9.
Once again, the signal component is shown in red and the background component in green. In addition, the tag components show the oscillating behaviour of this decay channel: the contribution from the unmixed decays $B_s^0 \to D_s^- \pi^+$ and $\bar{B}_s^0 \to D_s^+ \pi^-$ is depicted in black while the contribution from the mixed decays $B_s^0 \to \bar{B}_s^0 \to D_s^+ \pi^-$ and $\bar{B}_s^0 \to B_s^0 \to D_s^- \pi^+$, in purple. The parameters defining the distributions of signal and background are listed in the figure.

The complete signal distribution is therefore represented by the pdf:

$$PDF_{Sig}^{tot}\left(mass, time\right) = PDF_{mass} \times PDF_{Sig}\left(t, q_t, q_f|\omega\right) \tag{7.7}$$

where:

$$PDF_{mass} = f \cdot G_1\left(m, m_{B_s}, \sigma_1\right) + (1-f) \cdot G_2\left(m, m_{B_s}, \sigma_2\right) \tag{7.8}$$

describes the invariant mass distribution, and

$$PDF_{Sig}\left(t, q_t, q_f|\omega\right) = e^{-\frac{t}{\tau}}\left[\cosh\left(\frac{\Delta\Gamma_s t}{2}\right) + (1-2\omega)\cos\left(\Delta m_s t\right)\right] \tag{7.9}$$

corresponds to the time distribution described given in Eq. 6.4. The background component of the total distribution instead is described by the pdf:

$$\begin{array}{c} PDF_{Bkg}\left(m,t\right) = \left[f \cdot e^{m\alpha} + (1-f) \cdot G_{\Lambda_b}\left(\mu,\sigma\right)\right] \cdot PDF_{Non-osc}^{Tagging}\left(q_t, q_f, \epsilon_{tag}, \eta\right) \\ \cdot e^{-\Gamma t} \otimes R\left(t-t', \sigma_t - \sigma_{t'}\right) \cdot P\left(\omega\right) \end{array} \tag{7.10}$$

where:

$$PDF_{Non-osc}^{Tagging}(t, q_t, \epsilon_{tag}, \omega) = \begin{cases} (1-\epsilon_{tag}) & \text{if } q_t = 0 \\ \omega \cdot \epsilon_{tag} & \text{if } q_t = +1 \\ (1-\omega)\cdot\epsilon_{tag} & \text{if } q_t = -1 \end{cases}$$

The total signal yield from this fit is equal to $N_{Sig} = 25,586 \pm 418$ events, obtained the mass window $m_{B_s} \in [5350, 5800]$ MeV/c^2. When the yield is corrected to account for the full mass window $m_{B_s} \in [5320, 5420]$ MeV/c^2 the number of signal events becomes $N_{Sig} = 28,159 \pm 460$, in agreement with the sFit results $N_{Sig}^{sFit} = 27,965 \pm 196$. In addition, the frequency of the $B_s^0 - \bar{B}_s^0$ oscillations Δm_s is found to be equal to:

$$\Delta m_s = 17.766 \pm 0.028\ ps^{-1} (\text{only stat. error})$$

in good agreement with the recent measurement of Δm_s at LHCb [30]:

$$\Delta m_s = 17.768 \pm 0.023\ ps^{-1}$$

The fact that both the total $B_s^0 \to D_s^- \pi^+$ signal yield and the oscillation frequency Δm_s are in agreement with the latest LHCb measurements indicates that the distribution chosen to model the background is reliable.

The average reconstructed mistag, $\langle\eta\rangle$, which is obtained with this method is equal to:

$$\langle\eta\rangle = 0.3858$$

and the calibration parameters that are extracted with this method are:

$$\begin{array}{l} p_0^{bFit} = 0.400 \pm 0.010 \\ p_1^{bFit} = 1.173 \pm 0.130 \end{array}$$

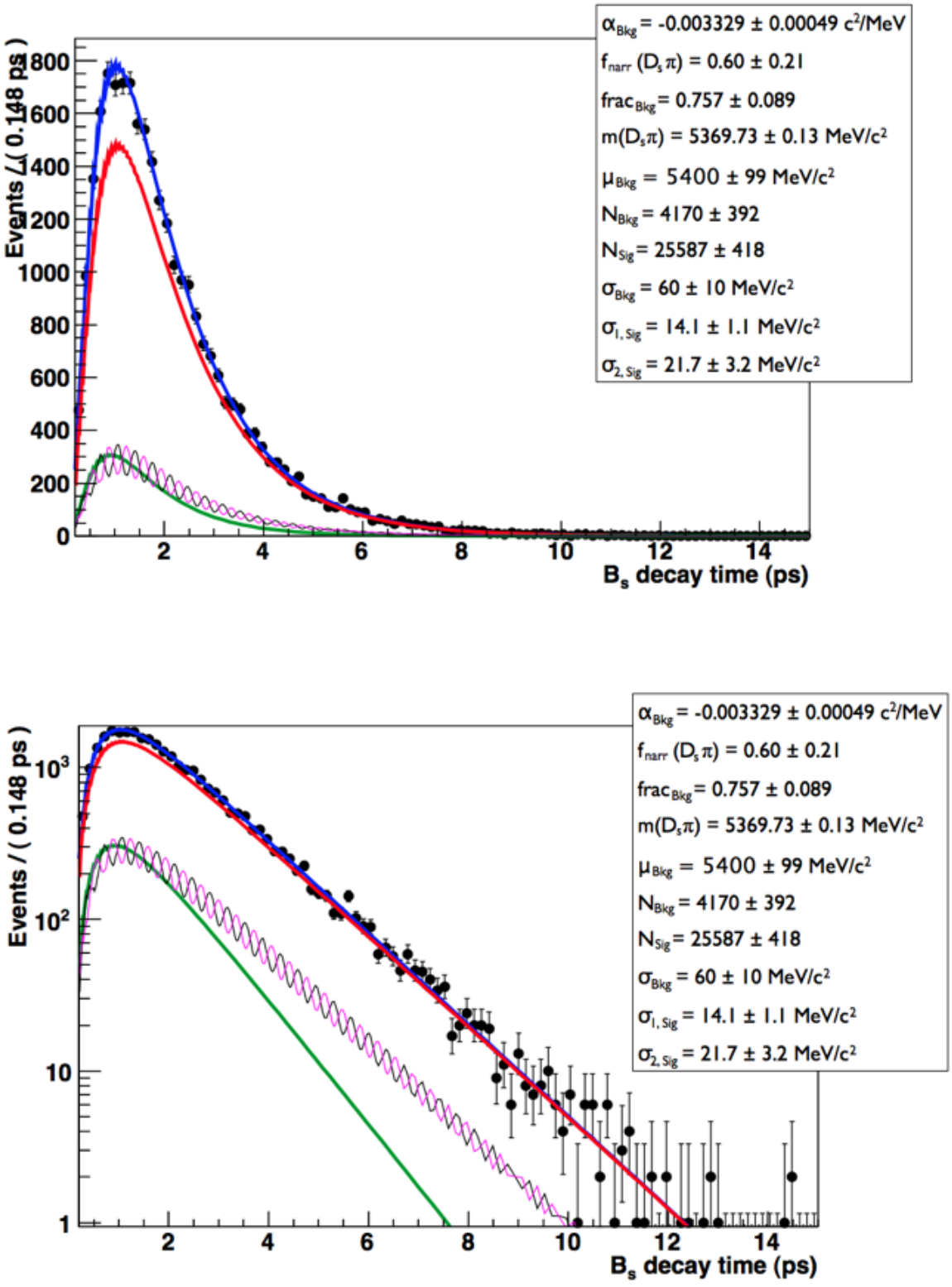

Figure 7.9: *Proper time distribution of $B_s^0 \to D_s^- \pi^+$ + c.c. candidates in linear scale (top) and logarithmic scale (bottom). The signal component is shown in red, the background component in green, the unmixed signal components $B_s^0 \to D_s^- \pi^+$ and $\bar{B}_s^0 \to D_s^+ \pi^-$ in black, and the contribution from the mixed decays $B_s^0 \to \bar{B}_s^0 \to D_s^+ \pi^-$ and $\bar{B}_s^0 \to B_s^0 \to D_s^- \pi^+$ in purple.*

to be compared to the sFit result:

$$p_0^{sFit} = 0.395 \pm 0.013$$
$$p_1^{sFit} = 1.369 \pm 0.152$$

showing good agreement on p_0 and a systematic effect on p_1 due to the different background treatment.

Since both the sFit extraction of the calibration parameters and the full fit performed in this section appear to be valid ways to extract p_0 and p_1, the final result we quote for the two is the arithmetic mean, and half the difference between the two methods assigned for the systematic uncertainty due to background modeling.

7.4.4 Results

The systematic errors on the parameters p_0 and p_1 that are due to a variation in the input parameters are summarized in Table 7.4. They are added in quadrature, in order to obtain the systematic uncertainties on the calibration parameters. Only the leading variations were taken into account, as these are the only systematic shifts to be significant for the measurement of p_0 and p_1. All variations on the parameter p_0 smaller than 10^{-3} and all those on the parameter p_1 smaller than 10^{-2} were neglected. In Table 7.6, the remaining variations that contribute to the systematic uncertainties on p_0 and p_1 are listed.

	$\lvert p_0 - p_0^+ \rvert$	$\lvert p_0 - p_0^- \rvert$	$\lvert p_1 - p_1^+ \rvert$	$\lvert p_1 - p_1^- \rvert$
SF_{res}	3.59×10^{-3}	5.51×10^{-3}	4.37×10^{-2}	6.88×10^{-2}
N_{cats}	1.49×10^{-3}	—	4.15×10^{-2}	2.33×10^{-2}
bFit	5.00×10^{-3}	—	0.196	—

Table 7.6: *Variations in the calibration parameters p_0 and p_1 which contribute to the systematic uncertainty on the measurement of these parameters. The shifts listed under the name bFit represent the difference between the results obtained from the sFit and those obtained from the fit including background events.*

As can be seen from the table, the leading variations on p_0 and p_1 are due to the systematic uncertainty on the scale factor applied to the proper time resolution (SF_{res}), to the variation in the number of mistag categories that are selected (N_{cats}) and to the background modeling in the data sample. It should be noted that the N_{cats} contribution includes a statistical contribution due to the limited statistics in the categories. The largest systematic comes from the treatment of the background.

The results obtained for the calibration parameters p_0 and p_1 are:

$$p_0 = 0.3975 \ \pm 0.0130 \text{ (stat.)} \pm 0.0059 \text{ (syst.)}$$
$$p_1 = 1.271 \ \pm 0.150 \text{ (stat.)} \pm 0.127 \text{ (syst.)}$$

The systematic and the statistical uncertainty on both calibration parameters p_0 and p_1 are comparable. These results are to be compared with the values of p_0 and p_1 determined by the LHCb tagging group using the high statistics sample of $B^+ \to J/\psi K^+$ data:

$$p_0 = 0.392 \pm 0.002$$
$$p_1 = 1.035 \pm 0.021$$

Since these results are obtained using independent data sets their difference should be compared to the total error. The difference (1.2σ) can also be due to a difference in the kinematics between $B^+ \to J/\psi K^+$ and $B_s^0 \to D_0^- \pi^+$ events. It is noted that the measurement will improve significantly once a larger statistics sample will be available for analysis.

Chapter 8

Sensitivity studies for the measurement of the angle γ

The results of the tagging are used to study the sensitivity of LHCb to measure the CKM angle γ using the time-dependent decay rates of $B_s^0 \to D_s^{\mp} K^{\pm}$ and $\bar{B}_s^0 \to D_s^{\pm} K^{\mp}$ events. This study will be based on the estimated statistics that will be collected in the second LHC run. A first time-dependent CP analysis on Run I is currently underway, but will not be sufficient for a precision measurement of γ using this decay mode.

Here we give an outlook for the expected sensitivity based on 4 years of data-taking and an integrated luminosity of 10 fb^{-1}, corresponding to the end of the end of the LHCb run II in 2017, as well as 50 fb^{-1}, expected to be collected after the LHCb upgrade. The observed decay time resolution in $B_s^0 \to D_s^{\mp} K^{\pm}$ events is about 40 fs, corresponding to a sensitivity reduction of 25% compared to a perfect resolution scenario (see Sec. 1.6). We determine the sensitivity of LHCb using the tagging performance as obtained in chapter 7.

In Sec. 8.1 a comparison is made for three methods:

- using a constant average wrong tag fraction for all events.
- assigning wrong tag values to event categories.
- assigning a reconstructed wrong tag value for each event.

In Sec. 8.2 the expected sensitivity on γ after the LHCb upgrade is presented.

8.1 Sensitivity on γ measured using $\langle\omega\rangle$, mistag categories and per-event ω

Fast MC studies are performed on $B_s^0 \to D_s^{\mp} K^{\pm}$ events in order to verify how much a precise calibration of the mistag probability ω improves the sensitivity on the measurement of the CKM angle γ. The angle γ, together with the physics parameters $|\lambda_f|$ and δ, can be extracted from the decay rate observables C_f, D_f, S_f, $\mathrm{D}_{\bar{f}}$ and $\mathrm{S}_{\bar{f}}$ as follows (see Sec. 1.5):

$$C_f = C_{\bar{f}} = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}$$

$$D_f = \frac{2|\lambda_f| \cdot \cos(\delta - (2\beta_s + \gamma))}{1+|\lambda_f|^2} \quad , \quad \mathrm{D}_{\bar{f}} = \frac{2|\bar{\lambda}_{\bar{f}}| \cdot \cos(\delta + (2\beta_s + \gamma))}{1+|\bar{\lambda}_{\bar{f}}|^2} \tag{8.1}$$

$$S_f = \frac{-2|\lambda_f| \cdot \sin(\delta - (2\beta_s + \gamma))}{1+|\lambda_f|^2} \quad , \quad \mathrm{S}_{\bar{f}} = \frac{-2|\bar{\lambda}_{\bar{f}}| \cdot \sin(\delta + (2\beta_s + \gamma))}{1+|\bar{\lambda}_{\bar{f}}|^2}$$

These coefficients are included in the full pdf used for these studies as described in Sec. 6.2. As can be seen from the above equations, for the decay $B_s^0 \to D_s^{\mp} K^{\pm}$ we have stated $C_f = C_{\bar{f}}$. This is motivated by two facts: first, because of CPT conservation, $|A_f| = |\bar{A}_{\bar{f}}|$ and $|A_{\bar{f}}| = |\bar{A}_f|$. Second, in the B_s system we expect no, or very small, CP violation in mixing; i.e. $|q/p| \simeq 1$. The combination of these two conditions implies $|\lambda_f| = |\bar{\lambda}_{\bar{f}}|$, which in turn leads to $C_f = C_{\bar{f}}$.

Three different scenarios are tested in order to determine how much the sensitivity on the measurement of the angle γ depends on the tagging performance. In the first scenario, an overall average mistag probability $\langle\omega\rangle$ is used in the fit performed to extract the parameters $|\lambda|$, δ and γ. In the second scenario several mistag categories are employed. After repeating this scenario several times, each time selecting a different number of mistag categories to see how the sensitivity on $|\lambda_f|$, δ and γ depends on the number of mistag categories used, the case of 7 mistag categories is chosen for further study. Lastly, the sensitivity on γ is determined by using the mistag probability that is assigned to each individual event during the reconstruction. In all cases we assume a perfect tagging calibration. We expect the use of mistag categories to significantly improve the precision on the measurement of the angle γ with respect to the value obtained using the average mistag probability $\langle\omega\rangle$. Intuitively, we subsequently expect the sensitivity on γ to perhaps marginally improve when the per-event mistag probability is used instead of the mistag categories.

The sensitivity study described in this section is performed neglecting backgrounds, i.e. only signal events are present in the analysed sample of simulated events. In order to simulate $B_s^0 \to D_s^{\mp} K^{\pm}$ data, the values of several parameters are fixed, such that the generated sample corresponds to the $B_s^0 \to D_s^{\mp} K^{\pm}$ data sample collected at LHCb during 2011 and 2012 and also to the sample to be collected during the LHCb run II that will take place from the start of 2015 through 2017. The $B_s^0 \to D_s^{\mp} K^{\pm}$ data yield for 2011 is equal to (1390 $\pm$ 98) events[1] [29], and twice as much data was collected in 2012 with respect to 2011. Including the period 2015-2017 we expect to collect about 10 fb^{-1} of integrated luminosity and run at a center of mass energy of 14 TeV, increasing the $b\bar{b}$ cross-section of 75% with respect to 2011 and leading to an estimated total of 9310 $B_s^0 \to D_s^{\mp} K^{\pm}$ recorded events for these years[2].
In this study, the total $B_s^0 \to D_s^{\mp} K^{\pm}$ yield is fixed to:

$$\mathrm{N}_{sig}^{(D_s K)} = 13965 \text{ events}$$

The values of the average mistag probability $\langle\omega\rangle$, tagging efficiency ϵ_{tag} and oscillation frequency Δm_s are extracted from our fit to $B_s^0 \to D_s^- \pi^+$ data. The values obtained for the tagging parameters, which are fixed in the fit to $B_s^0 \to D_s^{\mp} K^{\pm}$ data, are the average mistag probability $\langle\omega\rangle = 0.3833$ and the tagging efficiency $\epsilon_{tag} = 0.3927$. For the decay time resolution the fitted model of three Gaussians, as described in Sec. 6.3, with $\sigma_1 = 29.48$ fs, $\sigma_2 = 58.64$ fs and $\sigma_3 = 181.7$ fs is used. The B_s oscillation frequency Δm_s is taken to be $\Delta m_s = 17.76$ ps^{-1}, the decay width $\Gamma_s = 0.6580$ ps^{-1}, and the for the decay width difference the value $\Delta\Gamma_s = 0.116$ ps^{-1} is chosen.
Finally, the values of the ratio of interfering decay diagrams $|\lambda_f|$, the strong phase δ and the weak phase γ, which are used to simulate the fast MC data are set to:

$$|\lambda| = 0.311, \quad \delta = -35^\circ\,, \quad \gamma = 71.1^\circ.$$

Details on how the first two parameters were calculated can be found in chapter 1 [68], while the value for γ is taken from recent measurements [69].

[1] The available data sample in 2011 corresponds to 1.0 fb^{-1} of data from 7 TeV proton-proton collisions at the LHC.

[2] We consider the proper time resolution and tagging performance to remain unchanged from their values during the LHCb run I.

The proper time distribution for tagged events corresponding to 10 fb^{-1} of simulated data is shown in Figures 8.1 and 8.2. In the figure, decays of the B_s^0 and $\bar{B}_s^0$ to the final states $D_s^- K^+$ and $D_s^+ K^-$ are shown in the top and the bottom plot, respectively. From these distributions we expect that there is sensitivity to determine the time-dependent decay rates and therefore observe γ, but that the measurement will be limited by statistics. In Sec. 8.2 we therefore do a simulation to determine the sensitivity for the LHCb upgrade, with 50 fb^{-1} of simulated data.

To study the sensitivity, 2000 toy experiments are performed for each scenario. The results obtained for the physics parameters $|\lambda_f|$, δ and γ in the different scenarios are summarized in Table 8.1. The sensitivity on the measurement of each of these physics parameters is calculated by performing a Gaussian fit on the differences between the parameter's *true* value and its reconstructed value in many toy experiments and taking the width of the Gaussian. For example, the sensitivity on the measurement of the angle γ is given by $\sigma(\gamma_{true} - \gamma_{rec})$.

Scenario	$\sigma(\lvert\lambda_f\rvert_{true} - \lvert\lambda_f\rvert_{rec})$	$\sigma(\delta_{true} - \delta_{rec})$	$\sigma(\gamma_{true} - \gamma_{rec})$
average mistag	0.056 ± 0.001	(8.54 ± 0.11)°	(10.71 ± 0.17)°
7 mistag categories	0.049 ± 0.001	(7.33 ± 0.11)°	(9.85 ± 0.17)°
per-event mistag	0.047 ± 0.001	(7.22 ± 0.11)°	(9.74 ± 0.17)°

Table 8.1: *Measured sensitivity of the physics parameters $|\lambda_f|$, δ and γ for the different analysed scenarios, for 10 fb^{-1} of data.*

As expected, using an overall average value of the mistag probability ω gives a reduced sensitivity on the measurement of $|\lambda_f|$, δ and γ. The best precision on these parameters instead is obtained when the per-event mistag is employed, improving the sensitivity of about 10%. However, the difference with the mistag categories is only small, and the results obtained in these two scenarios are compatible within their uncertainty, confirming that dividing the data into categories having similar values of ω is an effective way of improving the sensitivity on γ.

In Fig. 8.3 the residuals and pull distributions of the parameters $|\lambda_f|$, δ and γ are shown for the scenario in which 7 mistag categories are used. The pull variable is defined as $(x - x_{true})/\sigma_x$, with $x = |\lambda|$, δ, or γ.
From the fitted mean pull distribution, a small bias in the measurement of $|\lambda_f|$ can be seen. However, no significant bias is present in the measurement of the angles δ and γ. In Table 8.2, the mean and width of the pull distribution of each parameter are summarized for the three mistag scenarios. The picture is confirmed in all cases: there is a small bias in $|\lambda_f|$ and no bias for δ and γ.

Scenario	Parameter	μ_{pull}	σ_{pull}
average mistag	$\lvert\lambda_f\rvert$	0.057 ± 0.023	1.006 ± 0.017
	δ	0.016 ± 0.023	1.002 ± 0.017
	γ	0.011 ± 0.024	1.038 ± 0.018
7 mistag cats	$\lvert\lambda_f\rvert$	0.070 ± 0.023	0.996 ± 0.017
	δ	0.013 ± 0.023	0.996 ± 0.017
	γ	0.011 ± 0.023	1.030 ± 0.017
per-event mistag	$\lvert\lambda_f\rvert$	0.091 ± 0.022	0.987 ± 0.017
	δ	0.015 ± 0.023	0.993 ± 0.017
	γ	0.007 ± 0.024	1.034 ± 0.018

Table 8.2: *Mean (μ_{pull}) and width (σ_{pull}) of the pull distributions of $|\lambda_f|$, δ and γ for the different analysed scenarios.*

The angles δ and γ are correlated with the value of $|\lambda_f|$ (see Eq. 8.1 for the CP asymmetry

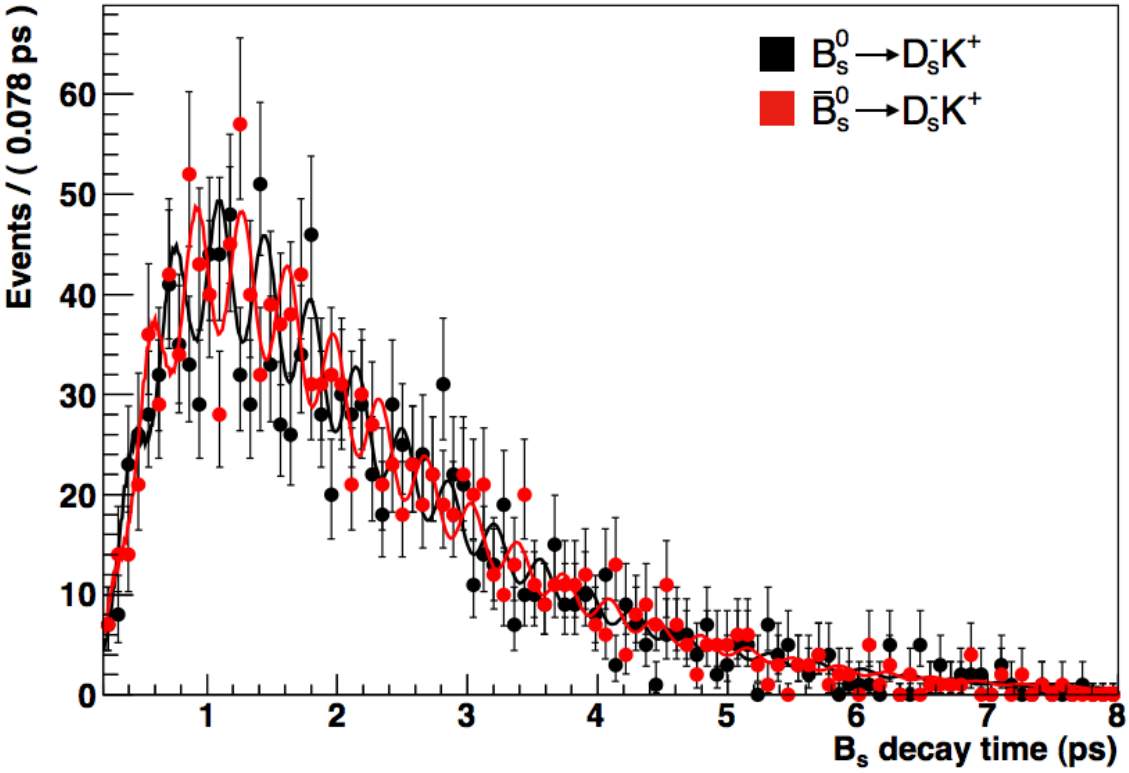

Figure 8.1: *Proper time distribution for tagged* $B_s^0 \to D_s^- K^+$ *and* $\bar{B}_s^0 \to D_s^- K^+$ *events obtained using 10* fb^{-1} *of simulated data.*

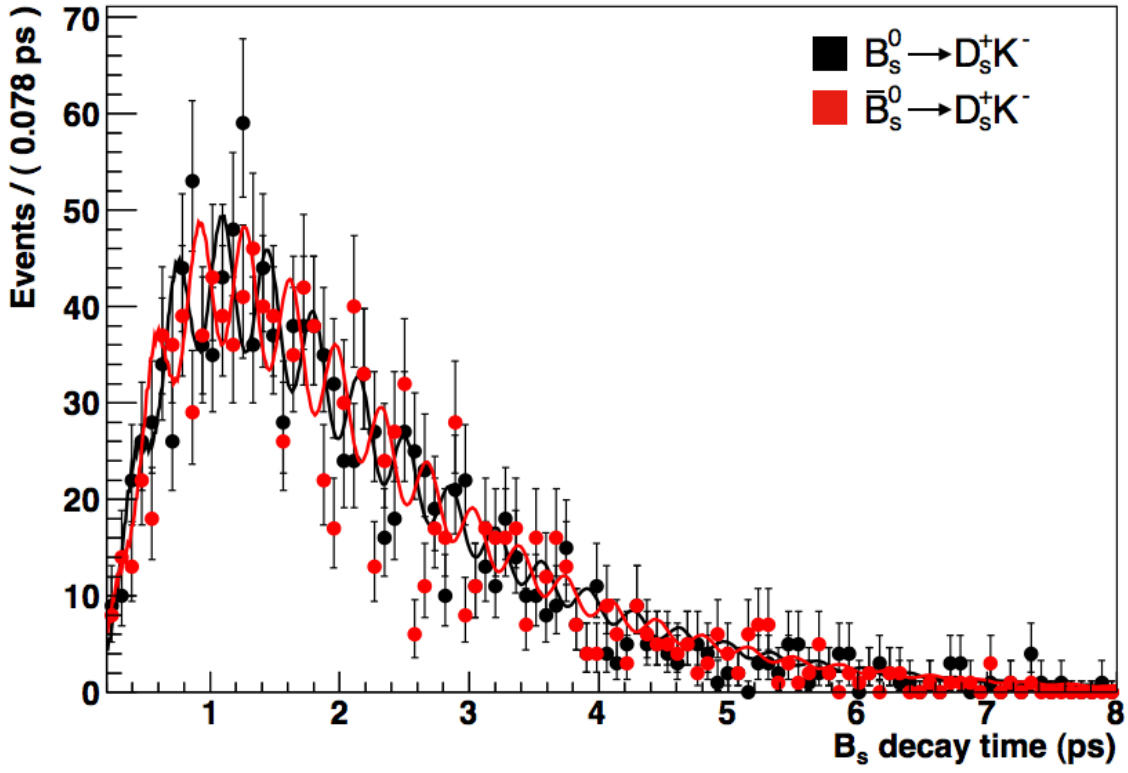

Figure 8.2: *Proper time distribution for tagged* $B_s^0 \to D_s^+ K^-$ *and* $\bar{B}_s^0 \to D_s^+ K^-$ *events obtained using 10* fb^{-1} *of simulated data.*

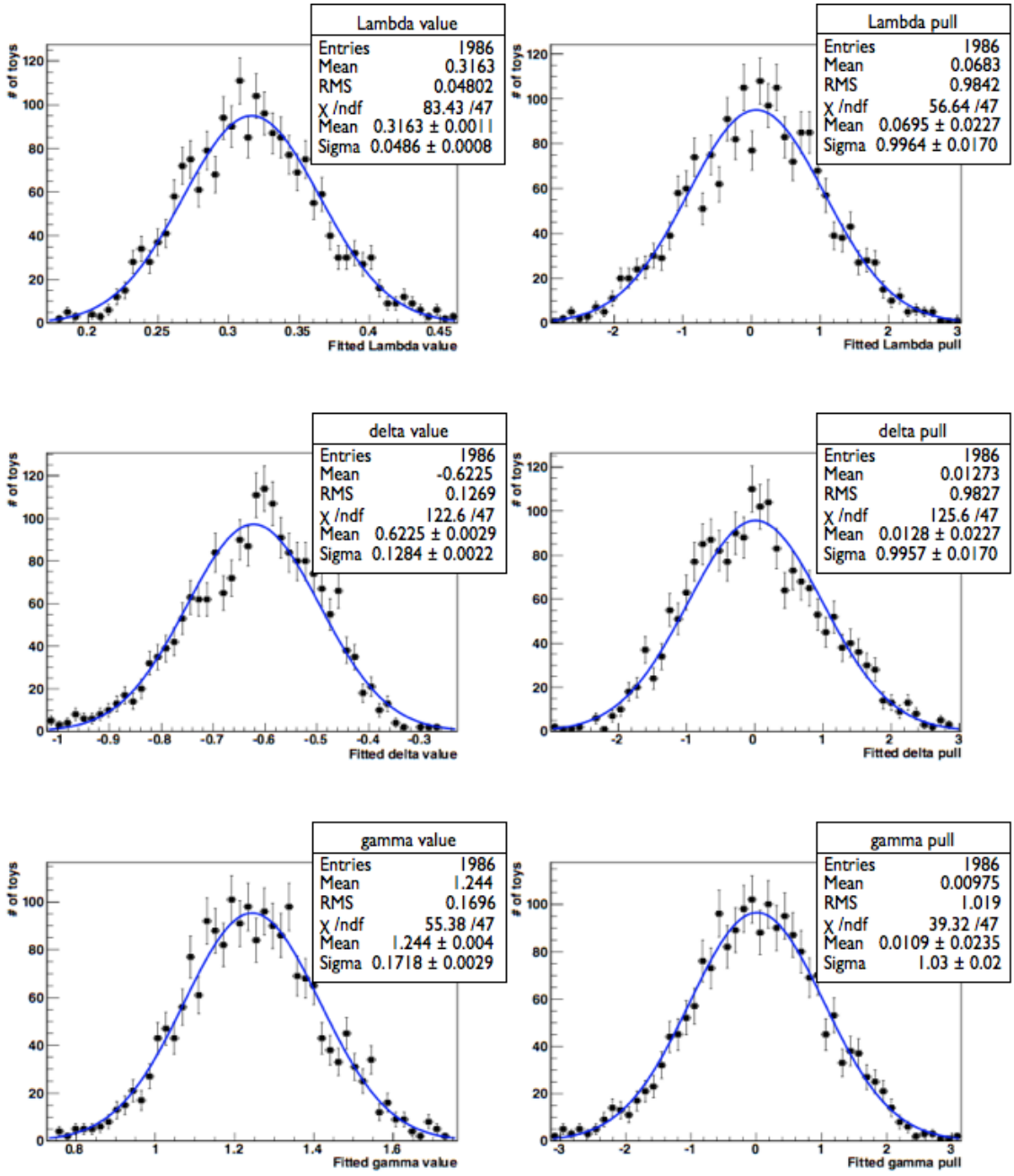

Figure 8.3: *Distributions of values and pulls of $|\lambda_f|$, δ and γ, for the scenario in which 7 mistag categories are selected, for a luminosity of 10 fb^{-1}. The values for δ and γ are plotted in radians. The input values of $|\lambda|$, δ and γ to the fit are $|\lambda| = 0.311$, $\delta = -35°$ (-0.611 radians), $\gamma = 71.1°$ (1.241 radians).*

observables C_f, D_f, $D_{\bar{f}}$, S_f and $S_{\bar{f}}$), implying that a bias in the value of $|\lambda_f|$ can also lead to a bias in the values of δ and γ. The magnitude of this bias depends on the degree of correlation between $|\lambda_f|$ and the angles δ and γ. However, in this analysis the correlation between these parameters is observed to be not problematic, as shown in Table 8.3.

	$\lvert\lambda_f\rvert$	δ	γ
$\lvert\lambda_f\rvert$	1.000	0.405	-0.097
δ	0.405	1.000	0.186
γ	-0.097	0.186	1.000

Table 8.3: *Correlation matrix for the parameters $|\lambda_f|$, δ and γ, for a typical toy experiment performed using 10 fb^{-1} of data.*

The results obtained for the sensitivity on the measurement of the angle γ show that using 10 fb^{-1} of data, i.e. at the end of 2014, a sensitivity of 10° on γ can be expected for the decay-time dependent measurement using $B_s^0 \to D_{\mp}K^{\pm}$ events. Compared to the uncertainties on the current combined measurement of $\gamma = (71.1^{+16.6}_{-15.7})^{\circ}$, the sensitivity of the time-dependent tagged analysis shows that a significant improvement can be made with this single, theoretically clean, measurement. In addition, a further improvement is possible if the mistag fraction could be reduced, and it has been shown that optimal performance is obtained using mistag categories instead of an average value of the mistag rate.

8.2 Outlook for the LHCb upgrade

During the first physics run 2011-2012, the LHCb detector has already performed better than foreseen. The number of interactions per beam crossing exceeded the design value of 0.4, such that the instantaneous luminosity was twice the design value. With the data recorded during these years, LHCb has produced several world best measurements of CP violation and mixing in the B meson sector. These results were obtained with the 3 fb^{-1} of pp data collected in 2011 and 2012, and are expected to further improve with the data that will be collected during the LHCb Run II, at the end of which we expect to have about 10 fb^{-1} of recorded data.

Although the results obtained by the LHCb spectrometer in the B_s^0-sector are already unique, an ambitious detector upgrade is proposed. The upgrade is scheduled to start data-taking in 2020 with an instantaneous luminosity of $2{\cdot}10^{33}$ cm^{-2}s^{-1}, five times the current value. In addition, the upgraded detector will have a full readout at 40 MHz allowing a more efficient trigger for hadronic final states. The study here is based on an integrated luminosity of 50 fb^{-1} and includes an estimated factor two improvement in trigger efficiency.

At a collision energy $\sqrt{s} = 14$ TeV, we expect the $b\bar{b}$ cross-section at LHCb to be 75% larger than that observed in 2011 and 2012. Therefore, we expect the total signal yield to be equal to:

$$\mathrm{N}_{sig}^{(D_sK)} = 232{,}750 \text{ events}$$

which corresponds to the number of $B_s^0 \to D_s^{\mp}K^{\pm}$ events expected to be observed in a data sample corresponding to approximately 50 fb^{-1}.

The upgraded LHCb detector is also designed such that the proper time resolution is expected to be about 15% better than the current one. Apart from an increased $B_s^0 \to D_s^{\mp}K^{\pm}$ signal yield and an improved proper time resolution, the tagging efficiency of the detector is also expected to improve, bringing the average mistag rate down from 38.3% to approximately 33.9% [70].

In Figures 8.4 and 8.5, the proper time distribution for tagged events is shown for 50 fb^{-1} of data. Once again, decays of the B_s^0 and $\bar{B}_s^0$ to the final states $D_s^-K^+$ and $D_s^+K^-$ are shown in the top and the bottom plot, respectively.

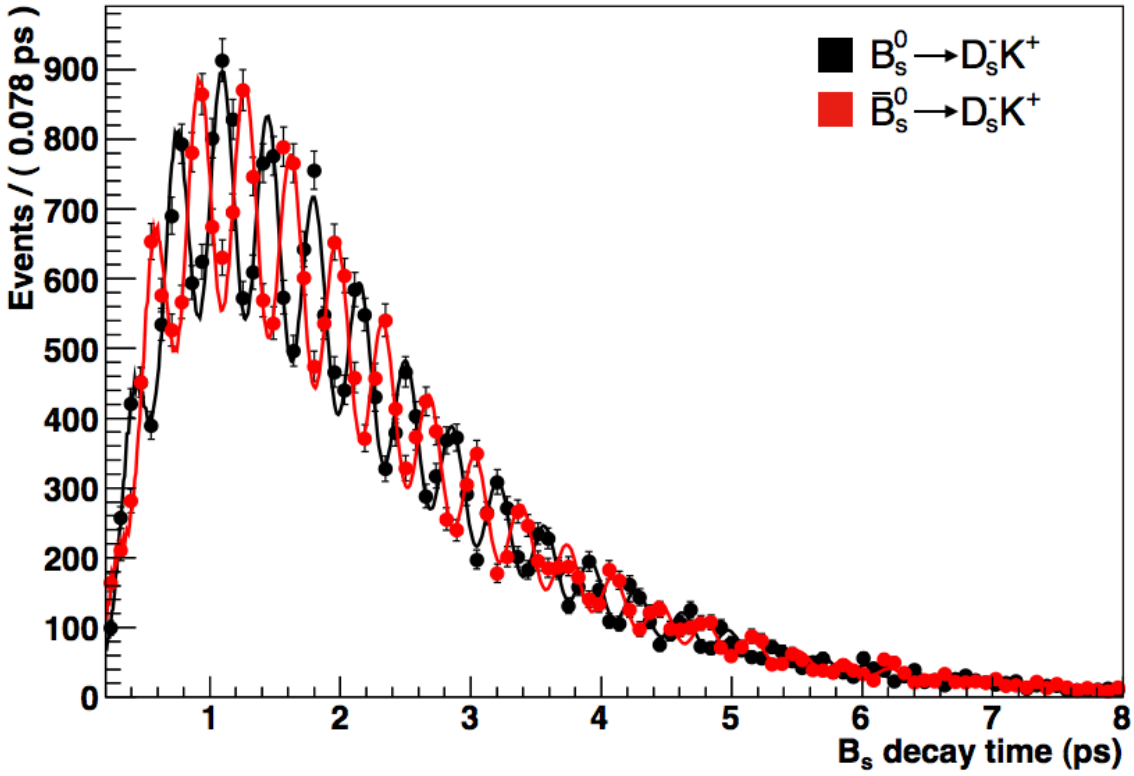

Figure 8.4: *Proper time distribution for tagged* $B_s^0 \to D_s^- K^+$ *and* $\bar{B}_s^0 \to D_s^- K^+$ *events obtained using 50* fb^{-1} *of data.*

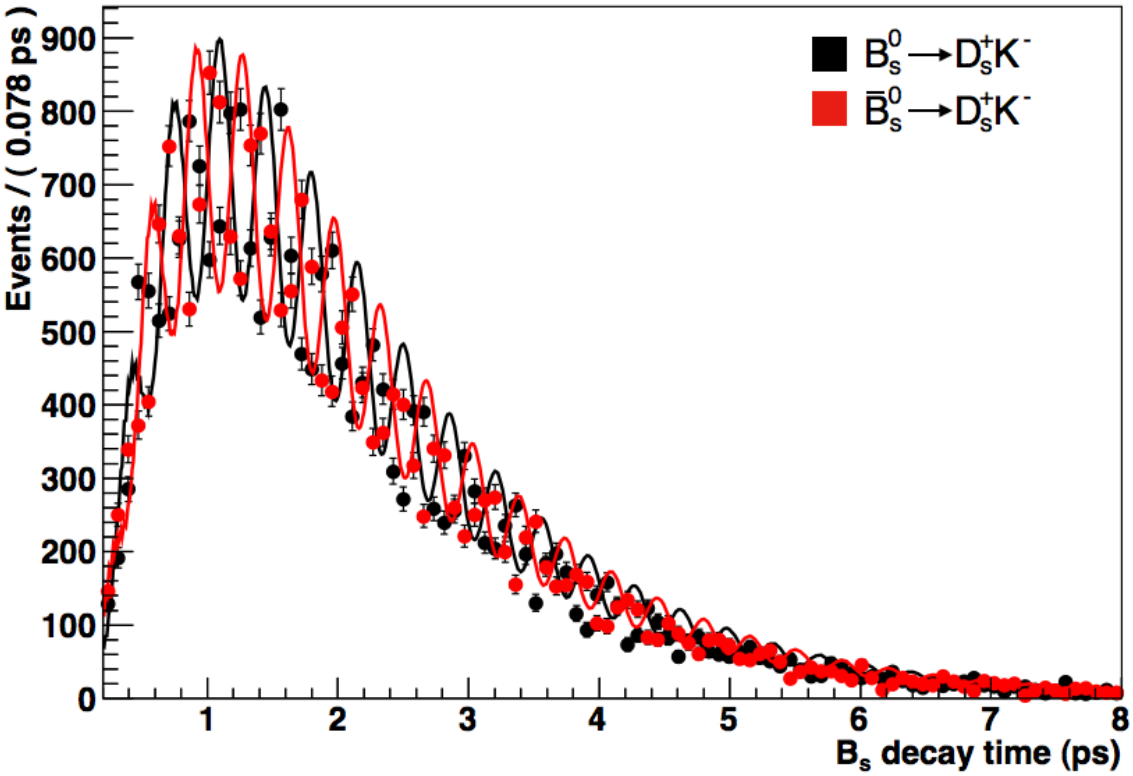

Figure 8.5: *Proper time distribution for tagged* $B_s^0 \to D_s^+ K^-$ *and* $\bar{B}_s^0 \to D_s^+ K^-$ *events obtained using 50* fb^{-1} *of data.*

It can be seen immediately that, thanks to the increase in statistics that will be available after the LHCb upgrade, the magnitude of the observed oscillations becomes larger and it becomes easier to disentangle the oscillations of the B_s^0 and $\bar{B}_s^0$ mesons.

An additional set of 2000 toy experiments was performed after reproducing the data-taking conditions after the LHCb upgrade, to verify how the sensitivity on the measurement of the CP angle γ will improve after this upgrade. Again, the toy experiments were performed in three scenarios: using an average value of the mistag rate, a per-event ω and dividing the sample into 7 mistag categories. For the three parameters, the distributions of values and pulls are depicted in Fig. 8.6.

The mean and width of the pulls for the parameters $|\lambda_f|$, δ and γ are listed in Table 8.4. A bias on the value of γ is observed, which could be due to the asymmetric shape of the likelihood around its minimum, due to the non-zero value of the strong phase δ. A more careful evaluation would be required for such high statistics data sample.

Scenario	Parameter	mean pull	width pull
average mistag	$\|\lambda_f\|$	-0.033 ± 0.022	0.982 ± 0.017
	δ	-0.034 ± 0.022	0.989 ± 0.017
	γ	0.108 ± 0.023	1.025 ± 0.018
7 mistag categories	$\|\lambda_f\|$	-0.045 ± 0.022	0.971 ± 0.017
	δ	-0.032 ± 0.022	0.989 ± 0.017
	γ	0.109 ± 0.023	1.021 ± 0.018
per-event mistag	$\|\lambda_f\|$	-0.030 ± 0.022	0.977 ± 0.017
	δ	-0.027 ± 0.023	0.992 ± 0.017
	γ	0.110 ± 0.023	1.022 ± 0.017

Table 8.4: *Mean and width of the pull distributions of $|\lambda_f|$, δ and γ, measured using a $B_s^0 \to D_s^{\mp} K^{\pm}$ sample collected with approximately 50 fb^{-1} of proton-proton collisions at the LHC.*

Comparing the pulls to the lower statistics study of 10 fb^{-1} shows that the bias in the determination of $|\lambda_f|$ has reduced, indicating a statistical effect. In fact, from Table 8.5 it can be seen that for 50 fb^{-1} the correlation between the parameters $|\lambda_f|$, δ and γ has reduced.

	$\|\lambda_f\|$	δ	γ
$\|\lambda_f\|$	1.000	0.189	-0.129
δ	0.189	1.000	-0.009
γ	-0.129	-0.009	1.000

Table 8.5: *Correlation matrix for the parameters $|\lambda_f|$, δ and γ, for a typical toy experiment performed using 50 fb^{-1} of data.*

In Table 8.6, the results obtained for the sensitivity on the parameters $|\lambda_f|$, δ and γ in the different scenarios of data collected before and after the LHCb upgrade are summarized.

The studies performed on simulated $B_s^0 \to D_s^{\mp} K^{\pm}$ data show a consistent picture that an improvement can be obtained using mistag categories or per-event wrong tag. The sensitivity on the angle γ measured using 50 fb^{-1} of simulated $B_s^0 \to D_s^{\mp} K^{\pm}$ data is equal to $(2.05 \pm 0.03)^{\circ}$, and represents an improvement of a factor of five compared to the sensitivity measured using 10 fb^{-1} of available data. A measurement of the angle γ at the few degree-precision level provides a stringent precision test for the CKM picture of CP violation.

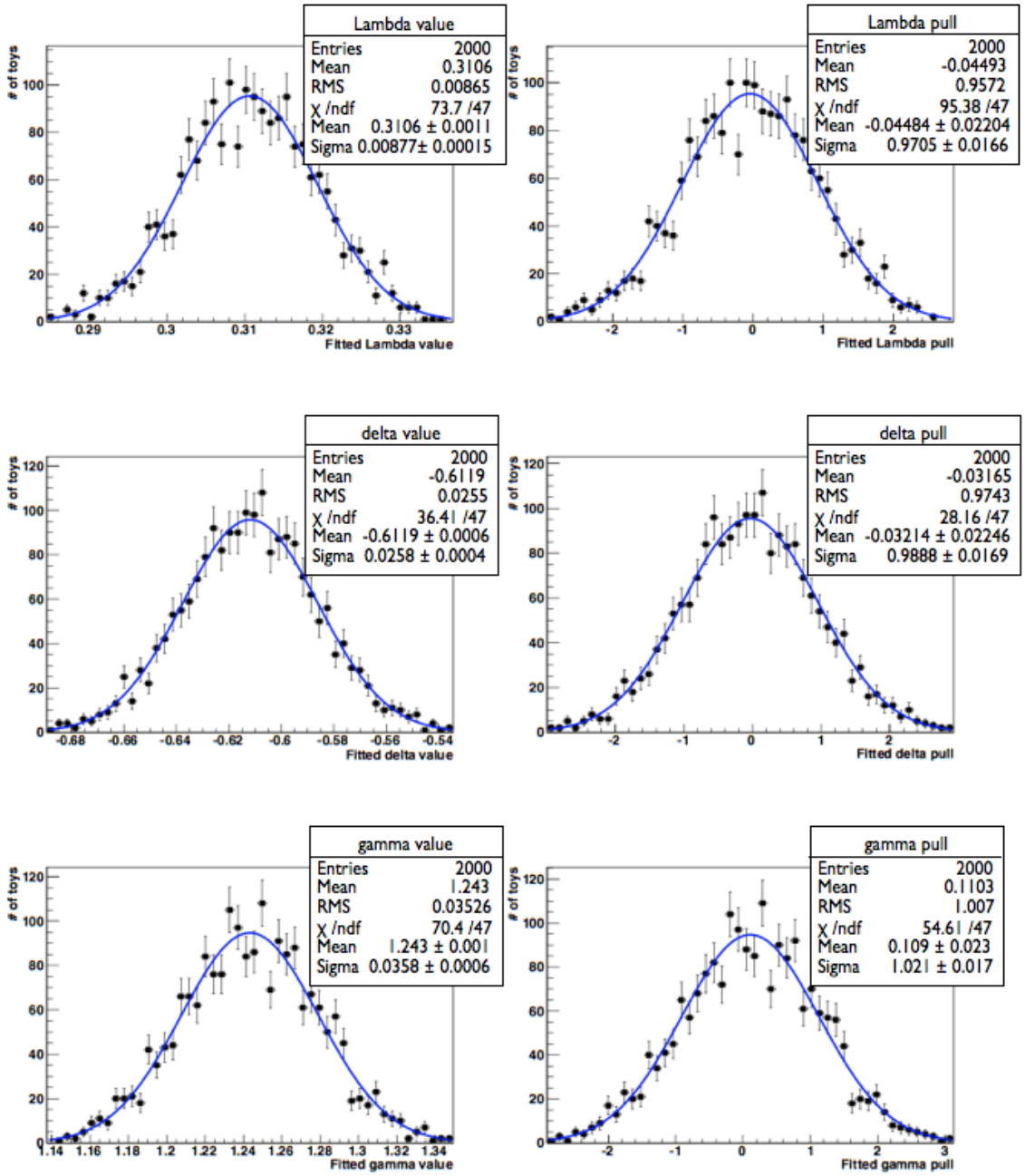

Figure 8.6: *Distributions of values and pulls of* $|\lambda_f|$, δ *and* γ, *for the scenario in which 7 mistag categories are selected, for a luminosity of 50* fb^{-1}. *The values for* δ *and* γ *are plotted in radians. The input values of* $|\lambda|$, δ *and* γ *to the fit are* $|\lambda| = 0.311$, $\delta = -35°$ *(*-0.611 *radians),* $\gamma = 71.1°$ *(1.241 radians).*

	$\sigma(\lvert\lambda_f\rvert_{rec} - \lvert\lambda_f\rvert_{true})$	$\sigma(\delta_{rec} - \delta true)$	$\sigma(\gamma_{rec} - \gamma_{true})$
10 fb^{-1}, $\langle\omega\rangle$	0.0564 ± 0.0010	(8.54 ± 0.11)°	(10.71 ± 0.17)°
10 fb^{-1}, 7 ω cats	0.0486 ± 0.0008	(7.33 ± 0.11)°	(9.85 ± 0.17)°
10 fb^{-1}, per-event ω	0.0474 ± 0.0008	(7.22 ± 0.11)°	(9.74 ± 0.17)°
50 fb^{-1}, $\langle\omega\rangle$	0.0097 ± 0.0002	(1.56 ± 0.03)°	(2.18 ± 0.03)°
50 fb^{-1}, 7 ω cats	0.0088 ± 0.0001	(1.48 ± 0.02)°	(2.05 ± 0.03)°
50 fb^{-1}, per-event ω	0.0087 ± 0.0001	(1.48 ± 0.02)°	(2.05 ± 0.03)°

Table 8.6: *Sensitivity on the parameters $|\lambda_f|$, δ and γ, measured in different scenarios using a 10 fb^{-1} and a 50 fb^{-1} sample of simulated $B_s^0 \to D_s^{\mp} K^{\pm}$ data.*

Measurements of the CKM angle γ have already been obtained via time-integrated methods using the charged $B^- \to D^{(*)0} K^{*-}$ and the neutral $B^0 \to D^{(*)0} K^{(*)0}$ decays [71, 72, 73]. However, the theoretically clean decay $B_s^0 \to D_s^{\mp} K^{\pm}$ is particularly suited to measure the weak phase γ, since the amplitudes of the decays $B_s^0 \to D_s^- K^+$ and $B_s^0 \to D_s^+ K^-$ are of the same order in the Wolfenstein parameter $\lambda - O(\lambda^3)$, meaning that interference effects in this decay channel are expected to be large. In addition, in the B_s^0 system the decay width difference $\Delta\Gamma_s$ is different from zero, increasing the sensitivity [66, 74].

The LHCb experiment provides an optimal environment to study γ, because of the large heavy flavour production cross-section and the excellent proper time resolution obtained with the LHCb detector. Care should be taken to obtain a bias-free measurement of the tagging dilution in the high pile-up environment. However, the large statistics that will be collected after the LHCb upgrade, combined with the above arguments, will allow for a precision test of the Standard Model via a measurement of the weak phase γ.

Appendix A

Raw data format and Event Info for the VELO

In the VELO detector, signals from the R and ϕ-sensors are sent via 60 m long cables to the TELL1 boards. To each sensor corresponds a TELL1 board, and the data that arrive at each TELL1 are processed by four Pre-Processing Field-Programmable-Gate Arrays (PP-FPGA). In each PP-FPGA, the information of the 4 Beetle chips is combined. A representation of the non-zero suppressed data for one PP-FPGA is shown in Fig. A.1.

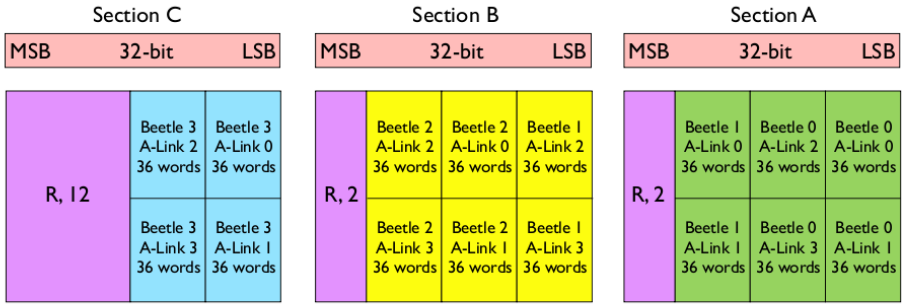

Figure A.1: *Packing of VELO non-zero suppressed 10-bit data in 32 bit words, for one PP-FPGA.*

The format of ZS data for a VELO event is shown in Fig. A.2 [75]. It is split into three parts: The user header (marked red), the cluster information block (marked yellow) and the ADC information block of all strips of the clusters (marked green).

- The user header consists of 32 bits of which 16 bits give the number of clusters (groups of 1, 2, 3 or 4 strips that have signal) in this event, 8 bits give the pipeline column number and 8 bits are spare.
- The cluster information block contains per cluster 1 spillover bit, 1 cluster size bit and the 14 bit cluster center position. The spillover bit indicates that the total charge corresponding to the sum of the ADC values for a cluster is above a given threshold. It is intended to distinguish between signal clusters and clusters caused by the signal spillovers from the previous event or noise. The cluster size bit distinguishes between clusters consisting of 1, 2, 3 or 4 strips. If the cluster has 1 or 2 strips this bit is set to 0, if there are 3 or 4 strips

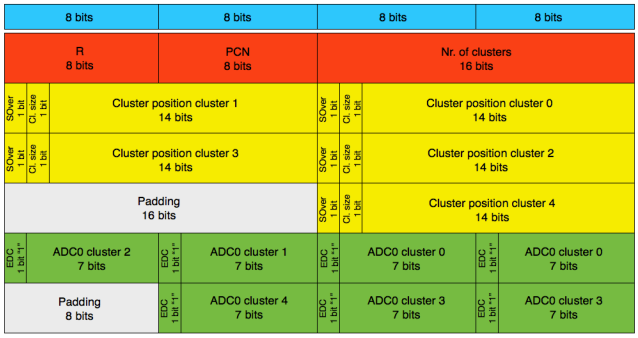

Figure A.2: *VELO ZS Data Format.*

in the cluster then this bit is set to 1. The last 14 bits give the cluster center in units of strips. Strips 0 to 2047 are encoded in 11 bits and the 3 least significant bits (LSB) give the inter-strip position. If the cluster size bit is 0 and the three LSBs of the cluster position are 000 then the cluster contains 1 strip. If the cluster size bit is 0 and at least one of the three LSBs of the cluster position is 1, it concerns a 2 strip cluster.

- The ADC information block contains the 7 bit ADC values of all strips in a cluster and the ADC end-of-cluster bit, which is set to 1 for the last ADC value of a cluster.

The "Event Info" for the VELO is the information collected from the PP-FPGAs and consists of 8 words [48]. The first three words are needed for event identification, synchronization check and to control processing options. When the data from the four PP-FPGAs is merged the three words are compared and if identical only the PCN is attached to the ZS data. If an error is detected all the 8 words are sent to the SyncLink-FPGA, so that an error bank containing complete error information can be generated. In Fig. A.3 the 8 words of Event Info for the VELO are shown.

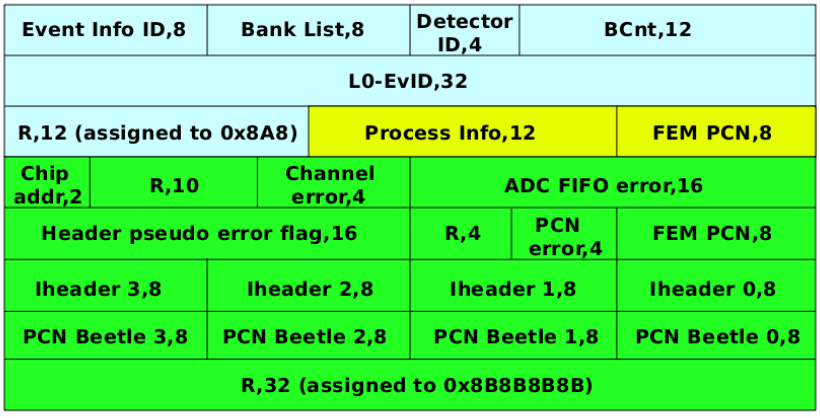

Figure A.3: *Eight words of the event info for the VELO are shown.*

The first word contains the event info ID, the bunch counter, detector ID (this corresponds to 0x1 for the VELO) and the bank list. The event info ID indicates what type of trigger was used

during data taking. The various bits of the event info, which are used for internal processing, define the following:

- Bit0-4: not used
- Bit5: ecs_trigger
- Bit6: data_generator_on
- Bit7: general_error_bit

In the bank list, the different bits define the following data banks:

- Bit0: info bank
- Bit1: cluster bank
- Bit2: ADC values bank
- Bit3: NZS bank
- Bit4: pedestal bank
- Bit5-7: not used

The cluster, ADC, NZS and pedestal banks only contain data words (no control words).

Words 2 and 3 of the Event Info contain the L0 Event ID, the FEM PCN and the Process Info. The FEM PCN is the PCN from the front-end emulator Beetle on the TELL1 and acts as reference, while the Process Info indicates what processing has been applied to the data. Words 4 and 5 contain error information. This includes the ADC FIFO error, the PCN error, the Header pseudo error flag and the channel error. These errors are discussed in detail in Ch. 3. The chip address indicates the Pre-processor from which the data come. Words 6 and 7 contain the 8-bit IHeader value and PCN value of each Beetle chip. The bits in the last word are reserved for later use.

Appendix B

Decoding of error banks

The decoding of the error banks is performed by an adaptive algorithm. This type of algorithm is needed because the structure of the error banks is not fixed; the number of words varies according to the number of errors present for one particular event. The decoding algorithm therefore has to adapt to the variable size of the error bank. In order to distinguish between error banks containing one, two or more errors, the decoder has to be able to recognize markers indicating the beginning of the different parts. Moreover, the decoder must be able to distinguish between the different PP-FPGAs in which the errors are observed. If, for example, errors are only present for PP-FPGA_0 and PP-FPGA_3, the words that contain information on the errors of PP-FPGA_3 will follow immediately after those of PP-FPGA_0. More detailed information regarding the decoding of error banks can be found in [76].

Bibliography

[1] F. Halzen, A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons Inc., 1984.

[2] The ATLAS Collaboration, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett. B* **716** no. 1 (2012) 1-29.

[3] The CMS Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett. B* **716** no. 1 (2012) 30-61.

[4] J.S. Bell, "Time Reversal in Field Theory", *Proc. R. Soc. Lond. A* **231** (1954) 479-495. Birmingham University Thesis.

[5] G. Lüders, "On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories", *Kongelige Danske Videnskabernes Selskab Matematisk-Fysiske Meddelelser* **28** n.5 (1954) 1-17.

[6] W. Pauli, *Niels Bohr and the Development of Physics*, McGraw-Hill, 1955.

[7] P.W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons". *Phys. Rev. Lett.* **13** (1964), 508-509.
F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (1964), 321-323.

[8] J.H. Christenson, J.W. Cronin, V.L Fitch, and R. Turlay, "Evidence for the 2π Decay of the K_2^0 Meson". *Phys. Rev. Lett.* **13** (1964) 138-140.

[9] The BaBar Collaboration, B. Aubert *et al.*, "Observation of CP violation in the B^0 meson system". *Phys. Rev. Lett.* **87** (2001) 091801.

[10] The Belle Collaboration, K. Abe *et al.*, "Observation of large CP violation in the neutral B meson system", *Phys. Rev. Lett.* **87** (2001) 091802.

[11] The D0 Collaboration, V.M. Abazov *et al.*, "Evidence for an anomalous like-sign dimuon charge asymmetry", *Phys. Rev. Lett.* **105** (2010) 081801.

[12] The CDF Collaboration, M. Morello, "CP and charge asymmetries at CDF", *J. Phys. Conf. Ser.* **110** (2008) 052040.

[13] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.* **10** (1963) 531533.

[14] M. Kobayashi and T. Maskawa, "CP violation in the Renormalizable Theory of Weak Interactions", *Progr. Theor. Phys.* **49** (1973) 652-657.

[15] W. Pauli, "Relativistic Field Theories of Elementary Particles", *Rev. Mod. Phys.* **13** (1941) 203-32.

[16] C.N. Yangs and R.L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance", *Physical Review* **96** (1954) 191-195.

[17] E. Noether, "Invariante Variationsprobleme", *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse* **1918** (1918) 235-257.

[18] S. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* **22** (1961) 579-588.

[19] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (1967) 1264-1266.

[20] A. Salam, "Weak and Electromagnetic Interactions", *Conf. Proc.* **C680519** (1968) 367-377.

[21] S. L. Glashow, J. Iliopoulos, L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry", *Phys. Rev. D*, **2** (1970) 1285-1292.

[22] The Particle Data Group Collaboration, J. Beringer *et al.*, "Review of Particle Physics (RPP)", *Phys. Rev. D* **86** (2012) 010001.

[23] C. Jarlskog, "Commutator of the quark mass matrices in the standard electroweak model and a measure of the maximal CP non conservation", *Phys. Rev. Lett.* (1985) 1039-1042.

[24] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa matrix", *Phys. Rev. Lett.* **51** (1983) 1945-1947.

[25] The Particle Data Group, K. Nakamura *et al.*, "Review of Particle Physics", *Journal of Physics G* **37** (2010) 075021.

[26] A.J. Buras, M.E. Lautenbacher, and G. Ostermaier, "Waiting for the top quark mass, $K^+ \to \pi^+$ neutrino anti-neutrino, $B^0_{(s)} - anti - B^0_{(s)}$ mixing and CP asymmetries in B decays", *Phys. Rev. D* **50** (1994) 3433-3446.

[27] J. Charles *et al.*, "CP violation and the CKM matrix: assessing the impact of the asymmetric B factories", *Eur. Phys. J. C* **41** (2005) 1-131.

[28] S. Ali, E. Rodrigues and M. Schiller, Private Communication, "Time-dependent CP Measurements: Master Equations and Conventions", *Tech. Rep.* (2011) LHCb-INT-2011-051.

[29] S. Ali *et al.*, "Measurement of time-dependent CP-violation observables in $B^0_s \to D^{\mp}_s K^{\pm}$", LHCb-ANA-2012-068 (2012).

[30] The LHCb Collaboration, R. Aaij *et al.*, "Precision measurement of the $B^0_s - \bar{B}^0_s$ oscillation frequency with the decay $B^0_s \to D^-_s \pi^+$", CERN-PH-EP-2013-054 (2013).

[31] M. Calvi *et al.*, "Flavour Tagging Algorithms and Performances in LHCb", CERN-LHCb-2007-058 (2007).

[32] L. Evans and P. Bryant, "LHC Machine", *Journal of Instrumentation* **3** no. 08, (2008) S08001.

[33] The LHCb Collaboration, R. Aaij *et al.*, "Measurement of the $B^{\pm}$ production cross-section in pp collisions at $\sqrt{s}$=7 TeV", *JHEP* **1204** (2012) 093.

[34] The LHCb Collaboration, P.R. Barbosa-Marinho, "LHCb VELO Technical Design Report", CERN-LHCC/2001-011 (2001).

[35] The LHCb Collaboration, A. A. Alves Jr. *et al.*, "The LHCb Detector at the LHC," *Journal of Instrumentation* **3** (2008) S08005.

[36] A. Kozlinskiy, *Outer tracker calibration and open charm production cross-section measurement at LHCb.* PhD thesis, Vrije Universiteit Amsterdam, 2013.

[37] N. Tuning, "Detailed Performance of the Outer Tracker at LHCb", *Journal of Instrumentation* **9** (2014) C01040.

[38] R. Fruhwirth, "Application of Kalman Filtering to Track and Vertex Fitting", *Nucl. Instrum. Methods Phys. Res. A*, **262** (1987) 444-450.
R.E. Kalman, "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35-45, 1960.

[39] J. van Tilburg, *Track simulation and reconstruction in LHCb* PhD thesis, Vrije Universiteit Amsterdam, 2005.
E. Bos, *Reconstruction of charged particles in the LHCb experiment* PhD thesis, Vrije Universiteit Amsterdam, 2010.
J. Amoraal, *Alignment of the LHCb Detector with Kalman Filter Fitted Tracks.* PhD thesis, Vrije Universiteit Amsterdam, 2011.

[40] S. Amato, J.R.T. De Mello-Neto, and C. Nunes, "Constrained and unconstrained vertex fitting tools for LHCb", LHCb-2004-038 (2004).

[41] X. Cid Vidal, "Muon Identification in the LHCb Experiment", *in the proceedings of the 45th Rencontres de Moriond on Electroweak Interactions and Unified Theories* (2010), `arXiv:1005.2585 [hep-ex]`.

[42] The LHCb RICH group, M. Adinal *et al.*, "Performance of the LHCb RICH detector at the LHC", CERN-LHCb-DP-2012-003 (2012).

[43] R.W. Forty and O. Schneider, "RICH pattern recognition", LHCb-98-040 (1998).

[44] O. Deschamps, F.P. Machefert, M.H. Schune, G. Pakhlova, and I. Belyaev, "Photon and neutral pion reconstruction", LHCb-2003-091 (2003).

[45] S. Oggero, *Beauty in the Crowd - Commissioning of the LHCb Pile-Up detector and first evidence of* $B_s^0 \to \mu^+\mu^-$, PhD thesis, Vrije Universiteit Amsterdam, 2013.

[46] P. Collins, Private Communication, "LHCb Central Exclusive Production Results and Prospects", LHCb-TALK-2013-199 (2013).

[47] G. Haefeli *et al.*, "The LHCb DAQ interface board TELL1", *Nucl. Instr. Methods Phys. Res. A* **560** no. 2, (2006) 494-502.

[48] O. Callot, M. Cattaneo, R. Jacobsson, B. Jost, P. Mato, N. Neufeld, "Raw-data Format", *EDMS* no. 565851/5 (2005).

[49] S. Klous, *LHCb: Vertex detector and read out chip, computing challenge and B(s)* $\to J/\psi\phi$ *analysis.* PhD thesis, Vrije Universiteit Amsterdam (2005).

[50] G. Haefeli, *Contribution to the development of the acquisition electronics for the LHCb experiment.* PhD thesis, École Polytechnique Fédérale de Lausanne (2004).

[51] S. Loechner and M. Schmelling, "The Beetle Reference Manual - chip version 1.3, 1.4 and 1.5", LHCb-2005-105 (2006).

[52] A. Gong, G. Haefeli, "VELO and ST non-zero suppressed bank data format", *EDMS* no. 692431 (2008).

[53] A. Gong, G. Haefeli, "VELO and ST error bank data format", *EDMS* no. 694818 (2009).

[54] The LHCb VELO group, M. Alexander *et al.*, "Performance of the LHCb VELO detector", *to be published in JINST.*

[55] The LHCb Collaboration, M.J. Charles *et al.*, "Performance of a wide pitch n-on-n silicon detector with floating strips", CERN-LHCb-2002-039 (2002).

[56] S. Ramo, "Currents induced by electron motion", *Proc. Ire.* **27** (1939) 584-585.

[57] Z. He, "Review of the Shockley-Ramo theorem and its application in semiconductor gamma-ray detectors", *Nucl. Instr. Methods Phys. Res. A* **463**, (2001) 250-267.

[58] The LHCb Collaboration, R. Aaij *et al.*, "Measurements of the branching fractions of the decays $B_s^0 \to D_s{}^{\mp}K^{\pm}$ and $B_s^0 \to D_s{}^{-}\pi^{+}$", *Journal of High Energy Physics* 1206 (2012) 115.

[59] The LHCb Collaboration, R. Antunes-Nobrega,"LHCb Trigger System Technical Design Report", CERN-LHCC-2003-031, LHCb-TDR (2003).

[60] M. Schiller, private communication, "Proper Time Resolution Studies in $B_s^0 \to D_s^- \pi^+$", LHCb-INT-2011-005 (2011).

[61] Y. Xie, "sFit: a method for background subtraction in maximum likelihood fit" (2009) arXiv:0905.0724 [physics.data-an].

[62] M. Pivk and F.R. Le Diberder, "SPlot: A Statistical tool to unfold data distributions" *Nucl. Instrum. Meth.* A555 (2005) 356-369.

[63] The LHCb Collaboration, R. Aaij *et al.*, "Opposite-side flavour tagging of B mesons at the LHCb experiment", *European Physical Journal C* **72** (2012) 2022.

[64] The LHCb Collaboration, R. Aaij *et al.*, "Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data", *Hadron Collider Physics Symposium* (2012) LHCb-ANA-2011-103.

[65] The LHCb Collaboration, R. Aaij *et al.*, "Performance of flavour tagging algorithms optimized for the analysis of $B_s^0 \to J/\psi\phi$", LHCb-CONF-2012-026 (2012).

[66] The LHCb Collaboration, R. Aaij *et al.*, "Tagged time-dependent angular analysis of $B_s^0 \to J\psi\phi$ decays at LHCb", CERN-LHCb-CONF-2012-002 (2012).

[67] The LHCb Collaboration, R. Aaij *et al.*, "Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data", LHCb-CONF-2012-033 (2012).

[68] K. De Bruyn, R. Fleischer, R. Knegjens, M. Merk, M. Schiller, N. Tuning, "Exploring $B_s \to D_s^{(*)\pm} K^\mp$ decays in the presence of sizable width difference $\Delta\Gamma_s$", *Nucl. Phys. B* (2012) 12663.

[69] The LHCb Collaboration, R. Aaij *et al.*, A measurement of γ from a combination of $B^\pm \to Dh^\pm$ analyses", LHCb-CONF-2012-032 (2012).

[70] The LHCb Collaboration, R. Aaij *et al.*, "Framework TDR for the LHCb Upgrade: Technical Design Report", CERN-LHCC-2012-007 (2012).

[71] I. Dunietz and R.G. Sachs, "Asymmetry between inclusive charmed and anticharmed modes in $B^0, \bar{B}^0$ decays as a measure of CP violation", *Phys. Rev. D* **37** (1988) 3186.

[72] R. Aleksan, I. Dunietz, and B. Kayser, "Determining the CP violating phase gamma", *Z. Phys. C* **54** (1992) 653-660.

[73] R. Fleischer, "New strategies to obtain insights into CP violation through $B_{(s)} \to D_{(s)}^\pm K^\mp$, $D_{(s)}^{(*)\pm} K^\mp$... and $B_{(d)} \to D^\pm \pi^\mp$, $D^{(*)\pm}\pi^\mp$, ... decays", *Nucl. Phys. B* **671** (2003) 459-482.

[74] A. Lenz and U. Nierste, "Numerical updates of lifetimes and mixing parameters of B mesons", *in the proceedings of the CKM workshop 2010 in Warwick* arXiv:1102.4274 [hep-ph].

[75] D. Eckstein, "VELO Raw Data Format and Strip Numbering", *EDMS* no. 637676 (2005).

[76] T. Szumlak, "VETRA- offline analysis and monitoring softaware platform for the LHCb Vertex Locator", *Journal of Physics: Conference Series* **219** (2010) 032058.

[77] The LHCb Collaboration, R. Aaij *et al.*, "First Evidence of Direct CP Violation in Charmless Two-Body Decays of B_s^0 Mesons", *Phys. Rev. Lett.* **108** (2012) 201601.

Web Sources

[78] The BaBar Collaboration, http://www-public.slac.stanford.edu/babar/.

[79] The Belle Collaboration, http://belle.kek.jp/.

[80] The CDF Collaboration, http://www-cdf.fnal.gov/.

[81] The D0 Collaboration, http://www-d0.fnal.gov/.

[82] The LHCb Collaboration, http://lhcb.web.cern.ch/.

[83] The CKMfitter Collaboration, http://ckmfitter.in2p3.fr.